

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 1, 2002, 14:43:43 ; Search time 2658 Seconds
(without alignments)
4532.943 Million cell updates/sec

Title: US-09-680-959-125
Perfect score: 414
Sequence: 1 ctaacgtacagaacagcttg.....gttnggaagcttgaccagt 414

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Se ad: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vl:*
29: em_htg_hum:*
30: em_htg_inv:*
31: em_htg_inv:*
32: em_htg_inv:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	384	92.8	2192	10	MMY15742	Y15742 Mus musculus
2	384	92.8	2329	10	BC030924	BC030924 Mus muscu
3	384	92.8	2428	10	MMBDYSTRO	AJ003007 Mus muscu
4	318.4	76.9	2221	9	BC016655	BC016655 Homo sapi
5	318.4	76.9	2244	9	HSDYSTRAO	Y15722 Homo sapien
6	318.4	76.9	2247	9	HSDYSTRAO	Y13712 Homo sapien
7	318.4	76.9	2407	9	AF022728	AF022728 Homo sapi
8	318.4	76.9	2573	9	HSDTMB2	Y13718 Homo sapien
9	316.8	76.5	2200	9	AK094494	AK094494 Homo sapi
10	202	48.8	380	10	MMU010206	AJ010206 Mus muscu
11	189.4	45.7	4615	5	FSCYRDYS	L05945 Torpedo cal
12	186.2	45.0	1593	10	AF143542	AF143542 Mus muscu
13	186.2	45.0	1883	10	MMDYSTM32	X55227 M.musculus
14	186.2	45.0	1969	10	AF143544	AF143544 Mus muscu
15	186.2	45.0	2607	10	MMDYSTM24	X55226 M.musculus
16	186.2	45.0	3144	10	AF143543	AF143543 Mus muscu
17	183	44.2	1644	9	HSA9668	AJ009668 Homo sapi
18	183	44.2	1707	9	HSU26742	U26742 Human dysr
19	183	44.2	1710	9	BC005300	BC005300 Homo sapi
20	183	44.2	2419	9	HSU26744	U26744 Human dysr
21	183	44.2	2518	9	HSU46744	U46744 Human dysr
22	183	44.2	2817	9	HSU46745	U46745 Human dysr
23	183	44.2	6066	9	HSM804598	AL833285 Homo sapi
24	178	43.0	204521	2	AC130150	AC130150 Rattus no
25	157.2	38.0	173893	9	AC010150	AC010150 Homo sapi
26	150.8	36.4	2818	3	AF277387	AF277387 Drosophi
27	121.8	29.4	1773	3	CEL131742	AJ131742 Caenorhab
28	100.8	24.3	351	10	MMDBNEX3	279789 M.musculus
29	99.8	24.1	525	9	HUMDTN03	U84531 Human dysr
30	99.8	24.1	15885	9	AC022601	AC022601 Homo sapi
31	99.8	24.1	175885	9	AC068506	AC068506 Homo sapi
32	99.8	24.1	178068	2	AF002411	AF002411 Homo sapi
33	94.2	22.8	62198	2	AC013832	AC013832 Drosophi
34	94.2	22.8	118100	2	AC017911	AC017911 Drosophi
35	94.2	22.8	171705	3	AC007476	AC007476 Drosophi
36	94.2	22.8	189557	3	AC007356	AC007356 Drosophi
37	94.2	22.8	233148	3	AE003822	AE003822 Drosophi
38	88.6	21.4	296	10	MMU010207	AJ010207 Mus muscu
39	85.4	20.6	146004	9	AC019144	AC019144 Homo sapi
40	72	17.4	306	11	DM49G4S	270900 D. melanoga
41	60.2	14.5	477	10	MMU010208	AJ010208 Mus muscu
42	44.6	10.8	256	10	MMDBNEX4	279790 M.musculus
43	44	10.6	105838	2	AC106483	AC106483 Rattus no
44	43.2	10.4	43956	3	AF098989	AF098989 Caenorhab
45	43.2	10.4	200509	2	AC006910	AC006910 Caenorhab

ALIGNMENTS

RESULT 1
MMY15742
LOCUS MMY15742 2192 bp mRNA linear ROD 29-MAY-1998
DEFINITION Mus musculus mRNA for dystrobrevin B, partial.
ACCESSION Y15742
VERSION Y15742.1 GI:3127923
KEYWORDS dystrobrevin B.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Puca, A.A.
REFERENCE 1 (bases 1 to 2192)
TITLE Direct Submission
JOURNAL Submitted (28-NOV-1997) A.A. Puca, T.I.G.E.M, Via Olgettina 58,

```
20132 Milano, ITALY
2 (bases 1 to 2192)
Pucca, A.A., Nigro, V., Piluso, G., Beisito, A., Sampaolo, S.,
Quaderi, N., Rossi, E., Di Iorio, G., Ballabio, A. and Franco, B.
Identification and characterization of a novel member of the
dystrobrevin gene family
FEBS Lett. 425 (1), 7-13 (1998)
98200066
9540997
Location/Qualifiers
1..2192
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="12"
1..2192
/feature="MDTN-B"
<1..>2192
/feature="MDTN-B"
/codon_start=1
/product="dystrobrevin B (MDTN-B)"
/protein_id="CAA75752.1"
/db_xref="GI:3127924"
/db_xref="SPTREMBL:O70585"
/translation="VKLTLTGKNLAAGEPRGSPAPAGKRRARIEEGNKRKTMWAE
KRLFTFEMAQNFVIRLSTYRTACKLRFVQKRNHLVDIWNMLEAPRDNGLNLTLDH
STEISVSRLETVISIYQLNRLPSTHQISVEQISISLLNFMVAAYSEGRGKLTVE
SVKMLATMCGGKMDKRLYIFSQMSDSNGLMFGKLDQFLKALKPLTAVPEGSPFG
YTEHAVRTCPQKKIMNLMFLDTMMADPPQCLVPLMLHRLAHVENVFFHVECSYC
HGESMVFYRQCQCHNYQLCQCFWRHSHAGSHQNMKEHSSWSPAKKLSHAIS
KSLGCVPSRPHVPEQPEKPLDLHLVPPRLTNMNDTVVSHMSGVPTTKRLQ
YSDQMPNLADLHAIASVYARLQHTVLDPSRLDEHRLIARYAAPAAEAGNMT
RPTDASFNDANKQORLIAELEKNREILQIOLRLEHEQASQPTPEKAQNPML
LAELRLRQKDELEQMSALQSRRELQVLEGLMKLKAQATGSPHSPHGGGRP
MMPVRSAGSTPTHGQDSLSGVGGVQEAFAQGTGRLNRLNLLVAADSTMTMS
LVKELHSGAEQETKREGLPEAHSCPSFCCTPGPSLQAARLTQPPRGAQPM
GSGEGLPEAAAGGPACTSIRIPPSWHGLISAGSIVQEMEKRNKALKK"
BASE COUNT 556 a 613 c 568 g 455 t
ORIGIN
Query Match 92.8%; Score 384; DB 10; Length 2192;
Best Local Similarity 98.5%; Pred. No. 3.5e-110;
Matches 406; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 3 AACGTACAGACAGCTTGCAGTTACCGATTGTACAGAGCGATGCAACCTTCATCTTG 62
DB 189 AACGTACAGACAGCTTGCAGTTA-CGATTTGTACAGAGCGATGCAACCTTCATCTTG 247
QY 63 TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGCGCTTAACACGCTGGACCACA 122
DB 248 TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGCGCTTAACACGCTGGACCACA 307
QY 123 GCACGGAGATCAGCGGTGTCCGCTTGAGACCGTCTATCTCGTCCATCTACTATCAGTTG 182
DB 308 GCACGGAGATCA-CGCGTGTCCGCTTGAGACCGTCTATCTCGTCCATCTACTATCAGTTG 366
QY 183 AACAGCGCTTCTCTTACTACAGATCAGCGTGGAGCAGTCCATCAGTCTCTCTACT 242
DB 367 AACAGCGCTTCTCTTACTACAGATCAGCGTGGAGCAGTCCATCAGTCTCTCTACT 426
QY 243 AATTTCATGTCGCGCTACGACAGTGGCGGAGGCGGAGCANGTTGACCGGTGTTTCAGNT 302
DB 427 AATTTCATGTCGCGCTACGACAGTGGCGGAGGCGGAGCANGTTGACCGGTGTTTCAGNT 486
QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTTC 362
DB 487 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTTC 546
QY 363 TCCAGATGTCAGATTCATTCAGGCTTAATGATGTTNGAAAGCTTGACCAAGT 414
DB 547 TCCAGATGTCAGATTCATTCAGGCTTAATGATGTTNGAAAGCTTGACCAAGT 598
RESULT 2
```

```
BC030924
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source
CDS
```

```
BC030924 2329 bp mRNA linear ROD 07-AUG-2002
Mus musculus, clone MGC:31669 IMAGE:4910166, mRNA, complete cds.
BC030924
BC030924.1 GI:21410283
MGC.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Prepared by: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 44 Row: k Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6681234.
Location/Qualifiers
1..2329
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:31669 IMAGE:4910166"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/feature="Vector: PCMV-SPORT6"
194..2002
/codon_start=1
/product="Unknown (protein for MGC:31669)"
/protein_id="AAH30924.1"
/db_xref="GI:21410284"
/translation="MIEGGNKRRTMAEKRLFIEMRAQNEVDVILSTYRTACKLRFV
QKRCNLHVDIWNMLEAFRDNGLNLTLDHSTEISVSRLETVISIYQLNKRUPSTHQI
SYEQISILLNFMVAAYDSEGRKLTVEFKAMLATMCGGKMDKRLYIFSQMSDSNG
LMFGKLDQFLKALKPLTAVPEGSPGYTEHAVRTCPQKKIMNLMFLDTPMADPP
PCLVPLMLHRLAHVENVFFHVECSYCHESMGFRYRCQCHNYQLCQCFWRGHA
SGAHSNQMKHEHSSWSPAKKLSHAISKSLGCVPSRPHVPEQPEKPLDLHLV
PPRLTNMNDTVVSHMSGVPTTKRLQYSDQMPNLADLHAIASVYARLQHTVRL
DPSRLDEHRLIARYAAPAAEAGNMTRPPTDASFNDANKQORLIAELEKNREI
LQEIQLRLEHEQASQPTPEKAQNPMLLAELRLRQKDELEQMSALQSRRELQV
LEGLMKLKAQATGSPHSPHGGGRPMPVRSAGSTPTHGQDSLSGVGGVQ
EFAQGTGRLNRLNLLVAADSTMTMSLVLKELHSG"
BASE COUNT 588 a 637 c 604 g 500 t
ORIGIN
Query Match 92.8%; Score 384; DB 10; Length 2329;
Best Local Similarity 98.5%; Pred. No. 3.6e-110;
Matches 406; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 3 AACGTACAGACAGCTTGCAGTTACCGATTGTACAGAGCGATGCAACCTTCATCTTG 62
DB 292 AACGTACAGACAGCTTGCAGTTA-CGATTTGTACAGAGCGATGCAACCTTCATCTTG 350
QY 63 TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGCGCTTAACACGCTGGACCACA 122
```



```

RESULT 6
HSDYSTRAO          2247 bp mRNA linear PRI 29-MAY-1998
LOCUS              Homo sapiens mRNA for brain dystrobrevin B.
DEFINITION         Y12712
ACCESSION          Y12712.1 GI:2765226
VERSION            dystrobrevin B.
KEYWORDS            Homo sapiens.
SOURCE             Homo sapiens
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS            Puca, A.A.
TITLE              Direct Submission
JOURNAL            Submitted (22-APR-1997) A.A. Puca, Seconda Università di Napoli,
INSTITUTE           Istituto di Patologia Generale, Larghetto S. Aniello a Caponapoli 2,
NAPOLI 80138, ITALY
2 (bases 1 to 2247)
Puca, A.A., Nigro, V., Piluso, G., Belisio, A., Sampaolo, S.,
Quaderi, N., Rossi, E., Di Iorio, G., Ballabio, A. and Franco, B.
Identification and characterization of a novel member of the
dystrobrevin gene family
FEBS Lett. 425 (1), 7-13 (1998)
98200066
9540997
FEATURES
Source            Location/Qualifiers
1..2247
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="p23"
/tissue_type="brain"
/dev_stage="adult"
184..1887
/codon_start=1
/product="dystrobrevin B"
/protein_id="CAA73249.1"
/db_xref="GI:2765227"
/db_xref="SPTREMBL:O43782"
/translation="MIEESGNKRRTMAEKROLFIEMRAQNFVIRLSTYRTACKLRFY
QKRCNLHLVDIWNMIEAFRDNLGNTLDHTTEISVSRLETIVISSIYQNLKRLPSTHQL
SVEQISILLNFMIAYDSDEGRGLTVFSVKAMLATMCGKMLDKLRYVFSOMSDSNG
LMIFSDFOLKVLKPLTAVFEGPSFGYTHSVTCFPQQRKIMLNFELDTMADPP
PCLVWLPLMHLAHVENVHPVECSYCRCSMGFRYCCQCHNYQIQCNCFWRGHA
GGPHSNQOMKEHSHWSKPAKLSHAIKSGUCVETPPPPVFPPEKPLDLAHIV
PPRLTNNDTMVSHMSGVPTTKSLVDSPLSRLEDEHRLIARYAARLAAAGNVRTP
PTDLNFNDANKQROLIAELNKNREILQETRLRLEHQASOPTPEKAQONPTLLA
ELRLRQRKDELRMSALQSRRLMVLQLELMKLKEEOKQAAQATGSPHTSPH
GGGRPMFVPRVSTAGSTPTTCHPQDLSLGVGDVQEAFAQAEAGEEERKQNGKDR
G"
BASE COUNT        591 a 606 c 612 g 438 t
ORIGIN
Query Match       76.9%; Score 318.4; DB 9; Length 2247;
Best Local Similarity 88.6%; Pred. No. 2.2e-89;
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;
QY 3 ACGGTACAGACAGCTTGCAGTTACCGATTGTTACAGAGCGATGCAACCTTCATCTTG 62
DB 282 AACTTACAGACAGCGCTGCAAAATTA-CGATTGTTACAAAACGATGCAACCTTCATCTTG 340
QY 63 TTGATATCTGGATATGATCGAGCTTCCGAGACAAATGCGCTTACACGCTGGACCACA 122
DB 341 TTGATATCTGGACATGATGATGAAGCTTCCGAGACAAATGCGCTTATACACCTGGACCATA 400
QY 123 GCACGGAGATCAGCGGTGTNCCGCTGGAGACCGTCTATCTGCTTACTATCATGTTG 182
DB 401 CCACCGAGATCA-GTGTGTCGCGCTCGGAACCTGCTCTCTCTTACTTACTATCATGTTG 459
QY 183 ACAAGCGCTTCTTCTTACTCACCAGATGAGCGTGGAGCAGTCCATCAGTCTCTCTACTC 242
DB 460 AACAGCGCTTCTTCTTACTCACCAGATGAGCGTGGAGCAGTCCATCAGTCTCTCTCTC 519

```

```

QY 243 AATTTTCATGTCGCGCTTACGACAGTGGAGCGCGGAGGANGTTGACCTGTTTTCAGNT 302
DB 520 AACTTTATGATGTCGTCATATGACAGTGGAGCGCGGAGGANGTTGACGTTATTTTCAGTT 579
QY 303 AAAGCTATGTTAGCAACACCATGCTGTGTGAAAAAATGCTGGACAAAATTTGAGATACATTTC 362
DB 580 AAAGCTATGTTAGCAACACCATGCTGTGTGAAAAAATGCTGGACAAAATTTGAGATACATTTC 639
QY 363 TCCAGATGTCAGATTCACATGCTTAATGCTTATGCTTNGGAAAGCTTGACCACT 414
DB 640 TCCAGATGTCAGATTCACATGCTTAATGCTTATGCTTNGGAAAGCTTGACCACT 691

RESULT 7
AF022728          2407 bp mRNA linear PRI 06-MAR-1998
LOCUS              Homo sapiens beta-dystrobrevin (BDTN) mRNA, complete cds.
DEFINITION         AF022728
ACCESSION          AF022728
VERSION            AF022728.1 GI:2935182
KEYWORDS            Homo sapiens.
SOURCE             Homo sapiens
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS            Peters, M.F., Jackson, K., Sadoulet-Puccio, H.M., Kunkel, L.M.,
ADAMS, M.E. and Froehner, S.C.
TITLE              beta-Dystrobrevin, a New Member of the Dystrophin Family:
IDENTIFICATION, CLONING, and Protein Associations
JOURNAL            Unpublished
REFERENCE           2 (bases 1 to 2407)
AUTHORS            Peters, M.F., Jackson, K., Sadoulet-Puccio, H.M., Kunkel, L.M.,
ADAMS, M.E. and Froehner, S.C.
TITLE              Direct Submission
JOURNAL            Submitted (03-SEP-1997) Physiology, UNC-CH, Rm 285 MSRB CB# 7545,
Chapel Hill, NC 27599, USA
FEATURES
Source            Location/Qualifiers
1..2407
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2p22-23"
1..2407
/gene="BDTN"
/gene="BDTN"
/feature="member of the dystrophin gene family"
/codon_start=1
/product="beta-dystrobrevin"
/protein_id="AAC05082.1"
/db_xref="GI:2935183"
/translation="MIEESGNKRRTMAEKROLFIEMRAQNFVIRLSTYRTACKLRFY
QKRCNLHLVDIWNMIEAFRDNLGNTLDHTTEISVSRLETIVISSIYQNLKRLPSTHQL
SVEQISILLNFMIAYDSDEGRGLTVFSVKAMLATMCGKMLDKLRYVFSOMSDSNG
LMIFSDFOLKVLKPLTAVFEGPSFGYTHSVTCFPQQRKIMLNFELDTMADPP
PCLVWLPLMHLAHVENVHPVECSYCRCSMGFRYCCQCHNYQIQCNCFWRGHA
GGPHSNQOMKEHSHWSKPAKLSHAIKSGUCVETPPPPVFPPEKPLDLAHIV
PPRLTNNDTMVSHMSGVPTTKSLVDSPLSRLEDEHRLIARYAARLAAAGNVRTP
PTDLNFNDANKQROLIAELNKNREILQETRLRLEHQASOPTPEKAQONPTLLA
ELRLRQRKDELRMSALQSRRLMVLQLELMKLKEEOKQAAQATGSPHTSPH
GGGRPMFVPRVSTAGSTPTTCHPQDLSLGVGDVQEAFAQAEAGEEERKQNGKDR
G"
BASE COUNT        630 a 654 c 649 g 474 t
ORIGIN
Query Match       76.9%; Score 318.4; DB 9; Length 2407;
Best Local Similarity 88.6%; Pred. No. 2.2e-89;
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;
QY 3 ACGGTACAGACAGCTTGCAGTTACCGATTGTTACAGAGCGATGCAACCTTCATCTTG 62
DB 282 AACTTACAGACAGCGCTGCAAAATTA-CGATTGTTACAAAACGATGCAACCTTCATCTTG 340
QY 63 TTGATATCTGGATATGATCGAGCTTCCGAGACAAATGCGCTTACACGCTGGACCACA 122
DB 341 TTGATATCTGGACATGATGATGAAGCTTCCGAGACAAATGCGCTTATACACCTGGACCATA 400
QY 123 GCACGGAGATCAGCGGTGTNCCGCTGGAGACCGTCTATCTGCTTACTATCATGTTG 182
DB 401 CCACCGAGATCA-GTGTGTCGCGCTCGGAACCTGCTCTCTTACTTACTATCATGTTG 459
QY 183 ACAAGCGCTTCTTCTTACTCACCAGATGAGCGTGGAGCAGTCCATCAGTCTCTCTACTC 242
DB 460 AACAGCGCTTCTTCTTACTCACCAGATGAGCGTGGAGCAGTCCATCAGTCTCTCTCTC 519

```

Db 260 AACATTACAGACAGCGCTGCAAAATTA-CGATTGTGACAAAACGATGCAACCTTCATCTTG 318
QY 63 TTGATATCTGGATATGATCGAGCTTTCCGAGACAAATGGCTTACACGCTGGACACA 122
Db 319 TTGATATCTGGACATGATGAGCTTCCGAGACAAATGGCTTATACACTGGACCA 378
QY 123 GCACGGAGATCAGGCGGTGTCGCGCTGGAGACCGTGCATCTGCTCCATCTACTATCAGTTG 182
Db 379 CCACCGAGATCA-GTGTGTCGCGCTCGAAACTGTCTCTCCATCTACTATCAGTTG 437
QY 183 AACACGGCGCTTCTCTCTACTACGACATCAGCGTGGAGCAGTCCATCAGTCTCTCTACT 242
Db 438 AACACGGCGCTTCTCTCTACTACCAAAATAGTGTGGAACAAATCTACTACGCGCTCTCTCT 497
QY 243 AATTTCATGTCGCGCTTCTCTCTACTACGACATCAGCGTGGAGCAGTGTGACCGTGTCTCTCAGNT 302
Db 498 AACTTTATGATGCTGCATATGACATGAGCGTGGAGCGGAGGCAAGTTGACGGTATTTTCAGTT 557
QY 303 AAAGCTATGTTAGCAACCAATGCTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 362
Db 558 AAAGCTATGTTAGCAACCAATGCTGTGGTGGAAAAATGCTGGACAAATTGAGATATGTTTTC 617
Db 363 TCCAGATGTCAGATCCCAATGCTTAATGATGTTNGGAAGCTTGACCACT 414
Db 618 TCCAGATGTCAGATCCCAATGCTTAATGATGTTNGGAAGCTTGACCACT 669

RESULT 8
HSDTNB2
LOCUS Homo sapiens mRNA for dystrobrevin B DTN-B2. PRI 27-MAY-1998
DEFINITION
ACCESSION Y15718
VERSION Y15718.1 GI:3133086
KEYWORDS DTN-B2; dystrobrevin B.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Puca,A.A., Nigro,V., Piluso,G., Belisito,A., Sampaolo,S.,
Quaderi,N., Rossi,E., Di Iorio,G., Ballabio,A. and Franco,B.
Identification and characterization of a novel member of the
dystrobrevin gene family
FEBS Lett. 425 (1), 7-13 (1998)
PUBMED 9540997
REFERENCE 2 (bases 1 to 2573)
AUTHORS Puca,A.A.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-1997) A.A. Puca, T.I.G.E.M. Via Olgettina 58,
MEDLINE 20132 Milano, ITALY
PUBMED
COMMENT Related sequence Y12712.
FEATURES
source
1..2573
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="p23"
/tissue_type="skeletal muscle"
/clone_lib="Clontech"
1..1677
/gene="DTN-B2"
1..1677
/gene="DTN-B2"
/codon_start=1
/product="dystrobrevin B DTN-B2"
/protein_id="CAAY533.1"
/db_xref="GI:3133087"
/db_xref="SPTREMBL:O60881"
/translation="MIEESGNKRTKMAEKRLQFLIEMRAQNFVIRLSTYRTACKLRVF
QKCNHLVDIWNMIEAFRONGNLTDHTTEISVRLTEYISVIYQLNKRLPSTHOI
SVQGSILLNFMIAVDSEGRKLTVESVKAMLATMCGGKMDLKLRYVFSOWSDSNG
LMIFSFKDFQLEKVLKLTAVFEGSPGFSYVTEHSVRCFPQQRKIMLNMFLLDTMMADPP

POCLVWLPMLHRLAHVENPVPBCSYCRCSMMGFYRCQOCHNYQLCONCFWRGHA
GGPSNQHMKEHSSWSPAKLSHAIKSLGCVPTREPPHVPPEQPEKDLDAHLIV
PPRLTNNMDTWVSHMSSGVPTPKRQYSDIDPSHLADEHALIASYVARLQHCARLV
DSPRLDEHLIARIARAAEAAGNVPTRPTDLSPFNDAKQORLIAELENCNREI
LQETQRLRLEHQASQPTPEKAQONPTLLAEIRLLRORKEDEORMASQSRRELMV
QLELMKLLKEEQAVSELNRNPPPSFARSFPHINYSPLPAHHL"
BASE COUNT 709 a 617 c 565 g 682 t
ORIGIN
Query Match 76.9%; Score 318.4; DB 9; Length 2573;
Best Local Similarity 88.6%; Pred. NO. 2.2e-89;
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;
QY 3 AACGTACAGAACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62
Db 99 AACTTACAGAACAGCTGCAAAATTA-CGATTGTGACAAAACGATGCAACCTTCATCTTG 157
QY 63 TTGATATCTGGATATGATCGAGCTTTCCGAGACAAATGGCTTACACGCTGGACACA 122
Db 158 TTGATATCTGGACATGATGAGCTTCCGAGACAAATGGCTTATACACTGGACCA 217
QY 123 GCACGGAGATCAGGCGGTGTCGCGCTGGAGACCGTGCATCTGCTCCATCTACTATCAGTTG 182
Db 218 CCACCGAGATCA-GTGTGTCGCGCTCGAAACTGTCTCTCCATCTACTATCAGTTG 276
QY 183 AACACGGCGCTTCTCTCTACTACGACATCAGCGTGGAGCAGTCCATCAGTCTCTCTACT 242
Db 277 AACACGGCGCTTCTCTCTACTACCAAAATAGTGTGGAACAAATCTACTACGCGCTCTCTCT 336
QY 243 AATTTCATGTCGCGCTTCTCTCTACTACGACATGAGCGCGGAGCGANGTTGACCGTGTCTCTCAGNT 302
Db 337 AACTTTATGATGCTGCATATGACATGAGCGTGGAGCGGAGGCAAGTTGACGGTATTTTCAGTT 396
QY 303 AAAGCTATGTTAGCAACCAATGCTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 362
Db 397 AAAGCTATGTTAGCAACCAATGCTGTGGTGGAAAAATGCTGGACAAATTGAGATGTTTTC 456
QY 363 TCCAGATGTCAGATCCCAATGCTTAATGATGTTNGGAAGCTTGACCACT 414
Db 457 TCCAGATGTCAGATCCCAATGCTTAATGATGTTNGGAAGCTTGACCACT 508

RESULT 9
AK094494 2200 bp mRNA linear PRI 15-JUL-2002
LOCUS Homo sapiens cDNA FLJ37175 fis, clone BRACE2028410, highly similar
DEFINITION to Homo sapiens beta-dystrobrevin (BDTN) mRNA.
ACCESSION AK094494
VERSION AK094494.1 GI:21753568
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens cerebellum cDNA to mRNA, clone_lib:BRACE2
clone:BRACE2028410.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Suzuki,Y., Sugano,S., Nagahara,K., Masubo,Y., Nagai,K.
and Isogai,T.
NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2200)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 252-0812, Japan
COMMENT (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers
1..2200
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2028410"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: pME18SFL3"
BASE COUNT 561 a 597 c 610 g 432 t
ORIGIN

Query Match 76.5%; Score 316.8; DB 9; Length 2200;
Best Local Similarity 88.3%; Pred. No. 6.9e-89;
Matches 364; Conservative 0; Mismatches 46; Indels 2; Gaps 2;
QY 3 AACGTACAGAACAGCTTGCAAGTTACCGATTGTACAGAGCGATGCAACCTTCATCTTG 62
DB 225 AACTTACAGAACAGCTGCAATTA-CGATTGTACAAAACGATGCAACCTTCATCTTG 283
QY 63 TTGATATCGGAATATGATCGAAGCTTCCGAGACAAATGCGCTTAACAGCGTGGACACA 122
DB 284 TTGATATCGGAACATGATGAAGCTTCCGAAACAAATGCGCTTAACACTGGACCA 343
QY 123 GCACGGAGATCAGGGGTGTCGCTGGAGAGCGTCATCTCGTCATCTACATCATGTTG 182
DB 344 CCACCGAGATCA-GTGTGTCGCGCTCGAACTGTCTCTCCATCTCTCCATCTATCAGTTG 402
-QY 183 AACAGCGCTTCTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 242
DB 403 AACAGCGCTTCTCTACTCACCAGATGATGAGCAATCTATCAGCGCTCCTCTC 462
QY 243 AATTCTATGTCGCGCTTACGACAGTGGAGCGGAGGAGTGGACGTTGTTTTCAGNT 302
DB 463 AACTTTATGATGCTGCATATGACAGTGGAGCGGAGGAGTGGACGTTTTCAGTT 522
QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGAAGAAATGCTGGACAAATGAGATACATTTTC 362
DB 523 AAAGCTATGTTAGCAACCATGTGTGGTGAAGAAATGCTGGACAAATGAGATATGTTTC 582
QY 363 TCCAGATGTCAGATTCGAATCCAAATGCTTATGATGTTNGGAAAGCTTGACCACT 414
D' 583 TCCAGATGTCAGATTCGAATCCAAATGCTTATGATGTTNGGAAAGCTTGACCACT 634

RESULT 10
LOCUS MMU010206 380 bp DNA linear ROD 30-NOV-1998
DEFINITION Mus musculus beta-dystrobrevin gene, exon 3.
ACCESSION AJ010206
VERSION AJ010206.1 GI:3702161
KEYWORDS beta-dystrobrevin; dtn-b gene.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 380)
AUTHORS Loh, N.Y.
TITLE Direct Submission
JOURNAL Submitted (05-AUG-1998) Loh N.Y., Genetics Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK
REFERENCE 2 (bases 1 to 380)
AUTHORS Loh, N.Y.; Ambrose, H.J.; Guay-Woodford, L.M.; DasGupta, S.; Nawrothki, R.A.; Blake, D.J. and Davies, K.E.
TITLE Genomic organization and refined mapping of the mouse beta-dystrobrevin gene
JOURNAL Mamm. Genome 9 (11), 857-862 (1998)

MEDLINE 99018217
PUBMED 9799833

FEATURES

Location/Qualifiers
1..380
/organism="Mus musculus"
/db_xref="taxon:10090"
84..297
/gene="dtn-b"
84..297
/gene="dtn-b"
/number=3
/usedin="AJ010204:dtn_CDS
/usedin="AJ010202:dtn_mRNA_A
/usedin="AJ010203:dtn_mRNA_B
BASE COUNT 84 a 109 c 70 g 117 t
ORIGIN

Query Match 48.8%; Score 202; DB 10; Length 380;
Best Local Similarity 99.1%; Pred. No. 1.2e-52;
Matches 213; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 54 TTCATCTGTTGATATCTCGAATATGATCGAAGCTTCCGAGACATGCGCTTAACACGC 113
DB 84 TTCATCTGTTGATATCTCGAATATGATCGAAGCTTCCGAGACATGCGCTTAACACGC 143
QY 114 TGGACACAGCAGGAGATCAGGCGTGTCCGCTGGAGACGTCATCTCGTCCATCTAC 173
DB 144 TGGACACAGCAGGAGATCA-GCGTGTCCGCTGGAGACGTCATCTCGTCCATCTAC 202
QY 174 TATCAGTTGAACAAGCGCTTCTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCACT 233
DB 203 TATCAGTTGAACAAGCGCTTCTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCACT 262
QY 234 CTCCTACTCAATTTTCATGTCGCGCGCTTACGACAG 268
DB 263 CTCCTACTCAATTTTCATGTCGCGCGCTTACGACAG 297

RESULT 11
LOCUS FSCITYRDYS 4615 bp mRNA linear VRT-22-JUN-1993
DEFINITION Torpedo californica 87 kDa postsynaptic membrane protein tyrosine kinase substrate dystrophin homolog mRNA.
ACCESSION L06945
VERSION L06945.1 GI:213244
KEYWORDS postsynaptic membrane protein.
SOURCE Torpedo californica (library: lamda gt11) adult electric organ cDNA to mRNA.
ORGANISM Torpedo californica
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes; Elasmobranchii; Squalae; Hypnosqualea; Pristiorajae; Batoidae; Torpediniformes; Torpedinoidei; Torpedinidae; Torpedo.
REFERENCE 1 (bases 1 to 4615)
AUTHORS Wagner, K.R., Cohen, J.B. and Haganir, R.L.
TITLE The 87k postsynaptic membrane protein from Torpedo is a protein-tyrosine kinase substrate homologous to dystrophin
JOURNAL Neuron 10 (3), 511-522 (1993)
MEDLINE 93213507
PUBMED 8461138

Location/Qualifiers
1..4615
/organism="Torpedo californica"
/db_xref="taxon:7787"
/tissue_type="electric organ"
/dev_stage="adult"
/tissue_lib="lamda gt11"
BASE COUNT 1507 a 836 c 893 g 1379 t
ORIGIN
Query Match 45.7%; Score 189.4; DB 5; Length 4615;
Best Local Similarity 69.1%; Pred. No. 1.8e-48;
Matches 284; Conservative 0; Mismatches 125; Indels 2; Gaps 2;

QY 4 ACGTACAGACAGCTTGCAGTACCGATTACCGATTTCACAGAGCGATCGAACCTTCATCTTGT 63
DB 235 ACCTACAGGACAGCATGTAGCT-TCGATTGTGAGAAAAATGCAATGTGCACTTGT 293
QY 64 TGATATCTGGAATATGATCGAAGCTTCCGAGACAAATGCCCTTAACACGCTGGACACAG 123
DB 294 AGATGTTGGAATGTATAGAGCATTCAGAGAAATGGTTAATACATGATCCGAA 353
QY 124 CACGAGATCAGCGGTGTCGCCCTGGAGACCGTCACTCGTCCATCTACTATCATAGTTGA 183
DB 354 CGCTGAACCTTA-GTGTGGCTGCTGGAAGCTATTATCTCCGCAATATCTTTACCAGCTGA 412
QY 184 ACAAGGCGCTCTCTACTACACAGATCAGCTGGAGAGTCCATCAGTCTCCTACTCA 243
DB 413 ACAAGAGATGCTTACACCCACAGATAACGTTGAGACGCTGCTAGTTTACTGCTGA 472
QY 244 ATTTTCATGTCGCCGCTACGACAGTGGAGCGGAGGANGTTGACCGTGTTCAGNTA 303
DB 473 ACTTCTGCTGCTGCTTATGATGAGAGGCGCATGGCAAAATATCAGTTTTTGTGTGA 532
QY 304 AGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTCT 363
DB 533 AAGTGTGTGTAGCAACATATGTCGAGGAAAGATTTCTGGATAAATGAGATATATTTCT 592
QY 364 CCAGATGTCAGATTCCTCAATGCTTAATGATCTTNGGAAAGCTTGACCACT 414
DB 593 CCCAAATATCAGACTCCCAATGGGTCATGATAAATGCAAGTTTGTATCAGT 643
RESULT 12
AF143542 1593 bp mRNA linear ROD 19-OCT-1999
LOCUS AF143542
DEFINITION Mus musculus alpha-dystrobrevin 3 mRNA, complete cds.
ACCESSION AF143542
VERSION AF143542.1 GI:4929244
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE
AUTHORS Enikg,R.E. and Maimone,M.M.
TITLE alpha-dystrobrevin isoform in muscle
JOURNAL Gene 238 (2), 479-488 (1999)
MEDLINE 20035753
PUBMED 10570976
REFERENCE 2 (bases 1 to 1593)
AUTHORS Maimone,M.M. and Enikg,R.E.
TITLE Direct Submission
JRNAL Submitted (16-APR-1999) Anatomy and Cell Biology, SUNY Health
Science Center at Syracuse, 750 East Adams St., Syracuse, NY 13210,
USA
FEATURES
source Location/Qualifiers
1. .1593
/organism="Mus musculus"
/strain="C3H"
/db_xref="taxon:10090"
/chromosome="18"
/map="between markers D18Mit62 and D18Mit23"
/clone="11.2"
/cell_line="BC3H1"
/tissue_type="muscle"
1. .244
245. .1360
/codon_start=1
/product="alpha-dystrobrevin 3"
/protein_id="AAD33913.1"
/db_xref="GI:4929245"
/translation="MIEDSGRRGNMFAEROLFAMRQADLDRLSTYRTACKLRFV
QKNCNHLVDIWNVLEALNNLDNIENLVARLEAVLSTFYQLNKRMPHTHQI
HVEQSLILLFLAADFEGHGKISFVAKMALATCGGKIMDKLRYIFSMISDSGG
VMVGRYDQFLREVLKLPATVAFEGSPFQTSARSCFSQOKKVTNLNGFLDLSMDPP

3'UTR polyA_signal 1570. .1575
BASE COUNT 435 a 390 c 346 g 422 t
ORIGIN
Query Match 45.0%; Score 186.2; DB 10; Length 1593;
Best Local Similarity 68.6%; Pred. No. 1.5e-47;
Matches 282; Conservative 0; Mismatches 127; Indels 2; Gaps 2;
QY 4 ACGTACAGACAGCTTGCAGTACCGATTACCGATTTCACAGAGCGATCGAACCTTCATCTTGT 63
DB 344 ACCATACAGACAGCATGCAAGCT-GAGATTGTGCGAGAAAGATGCAATTTGCACCTGT 402
QY 64 TGATATCTGGAATATGATCGAAGCTTCCGAGACAAATGCCCTTAACACGCTGGACACAG 123
DB 403 GGACATTTGGAACGTCATTGAAGCATTCGCGCAAAACGCTTGAATAACCTGGACCCCAA 462
QY 124 CACGAGATCAGCGGTGTCGCCCTGGAGACCGTCACTCGTCCATCTACTATCATAGTTGA 183
DB 463 CATAGAATCA-ACGTGGCCGCTGAGAGCGGTGCTCTCCACTATTTTAAQCACTCA 521
QY 184 ACAAGCGCTTCTCTACTACACAGATCAGCTGGAGCAGTCCATCAGTCTCCTACTCA 243
DB 532 ACAAGAGATGCCAACCACTCACCAAATCCACAGTGGAGCAGTCCATCAGTCTCCTGCTGA 581
QY 244 ATTTTCATGTCGCCGCTACGACAGTGGAGCGGAGGANGTTGACCGTGTTCAGNTA 303
DB 582 ACTTCTGCTGCGAGCTTTGACCGGGAAGCCATGGAAAAATCTCAGTATTTGCTGTCA 641
QY 304 AAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTCT 363
DB 642 AATGGCTTTGGCTACATTTGTGGAGGAAAGATCATCGACAAGTAAAGATATATTTCT 701
QY 364 CCAGATGTCAGATTCCTCAATGCTTAATGATCTTNGGAAAGCTTGACCACT 414
DB 702 CAATGATCTCTGACTCCAGTGGAGTGGTATATGGAAGATATGACCAAT 752
RESULT 13
MMDYSTM32 1883 bp mRNA linear ROD 24-APR-1996
LOCUS MMDYSTM32
DEFINITION M.musculus mRNA for dystrobrevin (clone m32).
ACCESSION X95227
VERSION X95227.1 GI:1246784
KEYWORDS 87 kDa protein; dystrobrevin.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Davies,K.E.
TITLE Isoform diversity of dystrobrevin, the murine 87-kDa postsynaptic
protein
JOURNAL J. Biol. Chem. 271 (13), 7802-7810 (1996)
MEDLINE 96205975
PUBMED 8631824
REFERENCE 2 (bases 1 to 1883)
AUTHORS Blake,D.J., Nawrotzki,R., Peters,M.F., Froehner,S.C. and
Davies,K.E.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-1996) D.J. Blake, Genetics Laboratory, Dep. of
Biochemistry, Univ. of Oxford, South Park Road, Oxford, Oxfordshire
OX1 3QU, UK
FEATURES
source Location/Qualifiers
1. 1883
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="m32"
/tissue_type="brain"
/dev_stage="adult"

```

CDS      176..1717
/codon_start=1
/product="dystrobrevin"
/protein_id="CAA64519.1"
/db_xref="GI:1246785"
/db_xref="MGI:106039"
/db_xref="SPTREMBL:Q61498"
/translation="MIEDSGKRGNTMAEROLFAPMRAQDLRLSTYRTACKLRFV
OKKCNLHVDIWNVTEALRENALNDPNIENVARLEAVLSTIFYOLNKRMPPTHQI
HVEQSIISLLNLLAFLAEPDPEGHGKISVFAVKMALATLCGGKIMDKLRYIFSMISDSSG
VMVGRYDOFLREVLLKLPATAVEGPGSEGYTEQSARSCFOOKKYVTLNGLFDTLMSDPP
PQCLVLLHRLANVENFHPVECSYCHSESMMGFYRCQCHNYQLCCQCFWRGHA
GGSHNQHMKEYTSWSPAKKLTNALSCSSREPLHPMFDPQPEKPLNLAHIV
DTWPPRYTSMNDTLFHSVPSSGSPFITRSMLESSNRLDEHRLIARYAARLAAESS
SOPPTQORASPDIFTIDANKOOROLIAELENKREILQEIQRLEVEHQASQPTPEK
AQONPTLLAELRLRQRKDELRMSALQESRRELVMQLEGLMKLLKEELKQGVSYV
PYCRS"
BASE COUNT      545 a 466 c 430 g 442 t
ORF --
Query Match      45.0%; Score 186.2; DB 10; Length 1883;
Best Local Similarity 68.6%; Pred. No. 1.6e-47;
Matches 282; Conservative 0; Mismatches 127; Indels 2; Gaps 2;
QY 4 ACCTACAGAACAGCTTGCAAGTTACCGATTGTTACAGAACGCGATGCAACCTTCATCTGT 63
Db 275 ACCTACAGAACAGCATGCAAGCT-GAGATTGTCAGAGAAATGCAATTTGCACCTGCT 333
QY 64 TGATATCTGGAATATGATGCAAGCTTCCGAGACATGCGCTTAACAGCTGGACACAG 123
Db 334 GGACATTTGGAACGTCATTGAAGCATTCGCGAAACGCTTTGAATAACCTGGACCCAAA 393
QY 124 CACGGAGATCAGCGGTGTCGCGTGGAGACGTCATCTCGTCCATCTACTATCATCTCA 183
Db 394 CATAGAACTCA-ACGTGGCGCGCTGGAGCGGTGCTCTCCACTATTTTACCAGCTCA 452
QY 184 ACAAGCGCTTCTCTACTACACAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTCA 243
Db 453 ACAAGAGGATGCCAACCACTCACCATAATCCAGTGGAGCAGTCCATCCTCTGCTCA 512
QY 244 ATTTCATGTCGCGCTAGCAGATGAGCGGGAGGCGAGGANTGTGACCGTGTTCAGNTA 303
Db 513 ACTTCCTGCTCGACGCTTTGACCGGAGGAGGCGATGGAATAATCTCAGTATTTGCTGCA 572
QY 304 AAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTCT 363
Db 573 AATGCTTTGCTAGCTATGTTGGAGGAAAGATCATGGAACAAGTTAAGATATATTTCT 632
QY 364 CCCAGATGTCAGATTCCTAATGCGTTAATGATGTTNGGAAAGCTTGACCACT 414
Db 633 CAATGATCTCTGACTCCAGTGGAGTGATGTTATGGAAGATATGACCAAT 683

RESULT 14
AF143544
LOCUS      Mus musculus alpha-dystrobrevin 2b mRNA linear ROD 19-OCT-1999
DEFINITION Mus musculus alpha-dystrobrevin 2b mRNA, complete cds.
ACCESSION AF143544
VERSION    AF143544.1 GI:4929248
KEYWORDS   Mus musculus.
SOURCE     Mus musculus.
ORGANISM   Mus musculus.
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      1 (bases 1 to 1969)
Enigk,R.E. and Maimone,M.M.
alpha-dystrobrevin isoform in muscle
JOURNAL    Differential expression and developmental regulation of a novel
MEDLINE    Gene 238 (2), 479-488 (1999)
PUBMED     20035753
REFERENCE  10570976
AUTHORS    2 (bases 1 to 1969)
Maimone,M.M. and Enigk,R.E.

TITLE Direct Submission
JOURNAL Submitted (16-APR-1999) Anatomy and Cell Biology, SUNY Health
Science Center at Syracuse, 750 East Adams St., Syracuse, NY 13210,
USA
FEATURES
source Location/Qualifiers
1..1969 /organism="Mus musculus"
/db_xref="BALB/c"
/db_xref="taxon:10090"
/chromosome="18"
/map="between markers D18Mit62 and D18Mit23"
/tissue_type="skeletal muscle"
1..190
5'UTR
CDS
191..1903
/codon_start=1
/product="alpha-dystrobrevin 2b"
/protein_id="AAD3915.1"
/db_xref="GI:4929249"
/translation="MIEDSGKRGNTMAEROLFAPMRAQDLRLSTYRTACKLRFV
OKKCNLHVDIWNVTEALRENALNDPNIENVARLEAVLSTIFYOLNKRMPPTHQI
HVEQSIISLLNLLAFLAEPDPEGHGKISVFAVKMALATLCGGKIMDKLRYIFSMISDSSG
VMVGRYDOFLREVLLKLPATAVEGPGSEGYTEQSARSCFOOKKYVTLNGLFDTLMSDPP
PQCLVLLHRLANVENFHPVECSYCHSESMMGFYRCQCHNYQLCCQCFWRGHA
GGSHNQHMKEYTSWSPAKKLTNALSCSSREPLHPMFDPQPEKPLNLAHIV
DTWPPRYTSMNDTLFHSVPSSGSPFITRSMLESSNRLDEHRLIARYAARLAAESS
SOPPTQORASPDIFTIDANKOOROLIAELENKREILQEIQRLEVEHQASQPTPEK
AQONPTLLAELRLRQRKDELRMSALQESRRELVMQLEGLMKLLKEELKQGVHIEPL
EERT"
3'UTR
polya_signal 1904..1969
BASE COUNT 584 a 477 c 459 g 449 t
ORIGIN
Query Match      45.0%; Score 186.2; DB 10; Length 1969;
Best Local Similarity 68.6%; Pred. No. 1.6e-47;
Matches 282; Conservative 0; Mismatches 127; Indels 2; Gaps 2;
QY 4 ACCTACAGAACAGCTTGCAAGTTACCGATTGTTACAGAACGCGATGCAACCTTCATCTGT 63
Db 290 ACCTACAGAACAGCATGCAAGCT-GAGATTGTCAGAGAAATGCAATTTGCACCTGCT 348
QY 64 TGATATCTGGAATATGATGCAAGCTTCCGAGACATGCGCTTAACAGCTGGACACAG 123
Db 349 GGACATTTGGAACGTCATTGAAGCATTCGCGAAACGCTTTGAATAACCTGGACCCAAA 408
QY 124 CACGGAGATCAGCGCTGTCGCGTGGAGACGCTCATCTCGTCCATCTACTATCAGTTGA 183
Db 409 CATAGAACTCA-ACGTGGCGCGCTGGAGCGGTGCTCTCCACTATTTTACCAGCTCA 467
QY 184 ACAAGCGCTTCTCTACTACACAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTCA 243
Db 468 ACAAGAGGATGCCAACCACTCACCATAATCCAGTGGAGGAGTCCATCAGTCTCCTGCTGA 527
QY 244 ATTTCATGTCGCGCTTACACAGATGAGCGCGGAGGCGAGGANGTTGACCGTGTTCAGNTA 303
Db 528 ACTTCCTGCTCGACGCTTTGACCGGAGGCGCATGGAATAATCTCAGTATTTGCTGTC 587
QY 304 AAGCTATGTTAGCAACCATGTGTGTGGAAAAATGCTGGACAAATGAGATACATTTCT 363
Db 588 AAATGGCTTTGGCTACATTTGTGTGGAGGAAGATCATGGAACAAGTTAAGATATATTTCT 647
QY 364 CCAGATGTCAGATTCCTAATGCTTAATGATGTTNGGAAAGCTTGACCACT 414
Db 648 CAATGATCTCTGACTCCAGTGGAGTGATGTTATGGAAGATATGACCAAT 698

RESULT 15
MMDYSTM24
LOCUS      MMDYSTM24 2607 bp mRNA linear ROD 24-APR-1996
DEFINITION M.musculus mRNA for dystrobrevin (clone m24).
ACCESSION X95226
VERSION    X95226.1 GI:1246782

```

KEYWORDS 87 kDa protein; dystrobrevin.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2607)
AUTHORS Blake,D.J., Nawrotzki,R., Peters,M.F., Froehner,S.C. and Davies,K.E.
TITLE Isoform diversity of dystrobrevin, the murine 87-kDa postsynaptic protein
JOURNAL J. Biol. Chem. 271 (13), 7802-7810 (1996)
MEDLINE 96205975
PubMed 8631824
REFERENCE 2 (bases 1 to 2607)
AUTHORS Blake,D.J.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-1996) D.J. Blake, Genetics Laboratory, Dep. of Biochemistry, Univ. of Oxford, South Park Road, Oxford, Oxfordshire OX1 3QU, UK
FEATURES
source 1..2607
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="m24"
 /tissue_type="brain"
 /dev_stage="adult"
 46..2112
 /codon_start=1
 /product="dystrobrevin"
 /protein_id="CAA64518.1"
 /db_xref="GI:1246783"
 /db_xref="MGD:MGI:106039"
 /db_xref="SPTREMBL:O61498"
 /translation="MIEDSGKRGNTMAERRQLFAEMRAQDLRIQLSTYRTACKLREV
 KKCNLHLVDINVLRENALNLDPIELNVLAEAVLSTIFYOLNKRMPPTHQI
 HVEQISLLNLFLLAEPDPEGHGKISVFAVKMALATLCGGKINDKLRVIFSMISDSGG
 VMYGRYDQFLREVLPKLTAVFEPSPGYTEQSNRSCFSQKKVTLNGFLDILMSDPP
 PQCLVPLHLRLANVENFVPCSYCHSESMGFYRCQCHNYQCCQCFWRGHA
 GGSHSNQHKQKTSWSPAKTKNALSKLSNREPLHMPDQPERPLNLAHIV
 PPRPVTSMDTLFSDVSPSSGPFITRSMLESSNRLDEHRLIARYAARLAASSSQ
 PTQORSADPISFTIDANKQORLIATELENKREILQIRLVEHEQASQTPKAAQO
 NPTLLARLLRQKDEQMSALQSRRELVMQLEGLMKLLKEELKQGTASSPR
 SSPSHTLSRPIPMPIRSASACPTTHTPQDSLTVGVGDVQEAFAQSSRNLRSDLLVA
 ADSTINTMSLVKELNLSVASETESESRFQFEDLAPSPYTSERAFIAQIHSRAP
 GYIHGGAASTHGDVDPDPPYTPQPDGNYENESVRLNELQLELYLQKLODEAY
 QVSLQG"
BASE COUNT 723 a 686 c 641 g 557 t
ORIGIN
Every Match 45.0%; Score 186.2; DB 10; Length 2607;
Local Similarity 68.6%; Pred. No. 1.7e-47;
Matches 282; Conservative 0; Mismatches 127; Indels 2; Gaps 2;
QY 4 ACGTACAGACAGCTTGCAGATTGCTACAGAGCGATGCAACCTTCATCTGT 63
DB 145 ACCTACAGACAGATGCAAGCT-GAGATTTGTGCGAAGAAATGCAATTTGCACCTGT 203
QY 64 TGATATCTGGATATGATCGAAGCTTTCAGACAAATGGCCCTTAACACGCTGGACACAG 123
DB 204 GGACATTTGGAACGTCATTGAAGCATTCGCGAANAACGCTTTGAATACTGGACCCAAA 263
QY 124 CAGGAGATCAGCGGTGTCGCGCTGGAGACGTCATCGTCCATCTACTATCAGTTGA 183
DB 264 CATAGAACTCA-ACGTGCGCGCTGGAGGGGGTCTCCACTATTTTACCAGCTCA 322
QY 184 ACAAGCGCTTCTTCTTACTCACCAGATCAGCGTGGAGCGATCCATCAGTCTCCTACTCA 243
DB 323 ACAAGAGATGCCAACCACTCACCAGATCCAGTGGAGCGATCCATCAGTCTCCTGCTGA 382
QY 244 ATTTTCATGGTCGCCCTACGACAGTGGAGGGCGCGAGCGANGTTGACCGTGTGTTTCAGNTA 303
DB 383 ACTTCTCTGCTCGACGCTTTGACCGGGAAGGCCATGGAAATCTCAGTATTGCTGTCA 442

QY 304 AAGCTATGTTAGCAACCATGCTGGTGGAAATGCTGGACAAATTGAGATACATTTTCT 363
DB 443 AAATGGGCTTTGGCTACATTTGTTGGAGGAAAGATCATGGACAAAGTTAAGATATATTTCT 502
QY 364 CCCAGATGTCAGATTCCTCAATGGCTTAATGATGTTNGGAAAGCTTTGACCAAGT 414
DB 503 CAATGATCTCTGACTCCACTGGAGTGATGCTATATGGAAGATATGACCAAT 553

Search completed: December 1, 2002, 17:08:47
Job time : 2665 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 14:38:13 ; Search time 256 Seconds
(without alignments)
3641.904.Million cell updates/sec

Title: US-09-680-959-125

Perfect score: 414
Sequence: 1 ctaacgtacagacagcttg.....gttnggaagcttgaccagt 414

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Se--hed: 2185239 seqs, 1125999159 residues

TOEAM number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	318.4	76.9	2247	ABK70286	Human lung cancer
2	183	44.2	2518	AAV18885	DNA encoding novel
3	181.4	43.8	2520	AAV18885	Human lung cancer
4	181	43.7	512	ABK70157	Drosophila melanog
5	150.8	36.4	3045	ABLI4199	Drosophila melanog
6	94.2	22.8	15987	ABLI4198	Human brain expres
7	81.4	19.7	472	AAQ60778	Human ovarian can
8	78.4	18.9	689	ABL87347	Human secreted pro
9	71.2	17.2	458	AAO1269	

10	39.2	9.5	5106	22	AAV31142	Human diagnostic a
11	38.2	9.2	13815	19	AAV18885	Mus musculus dystr
12	38.2	9.2	13815	24	ABK81960	cDNA encoding mous
13	38.2	9.2	13815	24	ABK199799	Mouse ischaemic co
14	38.2	9.2	19307	17	AAV21558	Shuttle vector pad
15	37.2	9.0	887	24	ABK81993	Human dystrophin c
16	37.2	9.0	1434	24	AAV37243	Human dystrophin r
17	37.2	9.0	1821	24	AAV37241	Human dystrophin r
18	37.2	9.0	2169	24	AAV37232	Human dystrophin r
19	37.2	9.0	3163	21	AAV48571	A rod shortened dy
20	37.2	9.0	3172	21	AAV48570	A rod shortened dy
21	37.2	9.0	3446	24	AAV37242	Human dystrophin m
22	37.2	9.0	3510	24	AAV37240	Human dystrophin m
23	37.2	9.0	3531	24	AAV37238	Human dystrophin m
24	37.2	9.0	3747	21	AAV48566	A rod shortened dy
25	37.2	9.0	3858	24	AAV37237	Human dystrophin m
26	37.2	9.0	3999	24	AAV37234	Human dystrophin m
27	37.2	9.0	4075	21	AAV48569	A rod shortened dy
28	37.2	9.0	4182	24	AAV37230	Human dystrophin m
29	37.2	9.0	4402	21	AAV48567	A rod shortened dy
30	37.2	9.0	4402	21	AAV48568	A rod shortened dy
31	37.2	9.0	4414	24	AAV37260	Adeno-associated v
32	37.2	9.0	4476	24	AAV37259	Adeno-associated v
33	37.2	9.0	4498	24	AAV37258	Adeno-associated v
34	37.2	9.0	4825	24	AAV37257	Adeno-associated v
35	37.2	9.0	4848	24	AAV37263	Adeno-associated v
36	37.2	9.0	4966	24	AAV37256	Adeno-associated v
37	37.2	9.0	4990	24	AAV37262	Adeno-associated v
38	37.2	9.0	5060	24	AAV37264	Adeno-associated v
39	37.2	9.0	5149	24	AAV37255	Adeno-associated v
40	37.2	9.0	5339	24	ABK81998	DNA encoding mini-
41	37.2	9.0	5417	24	ABK81997	DNA encoding mini-
42	37.2	9.0	5462	24	ABK81999	DNA encoding mini-
43	37.2	9.0	5952	22	AAV06794	Human dystrophin g
44	37.2	9.0	8689	24	ABK82000	DNA encoding mini-
45	37.2	9.0	11058	24	AAV37229	Human dystrophin p

ALIGNMENTS

RESULT 1
ABK70286
ID ABK70286 standard; cDNA; 2247 BP.
XX
AC ABK70286;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human lung cancer associated full length cDNA DMSM-52.
XX
KW Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.
XX
OS Homo sapiens.
XX
PN WO200224057-A2.
XX
PD 28-MAR-2002.
XX
PF 20-SEP-2001; 2001WO-USA2232.
XX
PR 22-SEP-2000; 2000US-234837P.
PR 10-OCT-2000; 2000US-239440P.
PR 29-JUN-2001; 2001US-301928P.
XX
(CORI-) CORIXA CORP.
XX
Benson DR, Mohamath R, Lodes MJ;
XX
WPI; 2002-372001/40.
XX
New tumour lung proteins and nucleic acids encoding the proteins, useful
as vaccines and for treating, preventing, diagnosing or monitoring lung

PT cancer
PS Claim 1; Page 160-161; 189pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising a sequence
CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),
CC or their fragments, homologues, variants or complements and their encoded
CC polypeptides. Also included are an expression vector comprising the
CC polynucleotide operably linked to an expression control sequence; a host
CC cell transformed or transfected with an expression vector of; an isolated
CC polypeptide; a method for detecting the presence of a cancer in a
CC patient; a fusion protein comprising at least the polypeptide; an
CC oligonucleotide that hybridises to the polynucleotide under moderately
CC stringent conditions; a method for stimulating and/or expanding T cells
CC specific for a tumour protein; an isolated T cell population comprising T
CC cells prepared from the method of above; a composition comprising a first
CC component consisting of carriers and immunostimulants, and a second
CC component selected from the polynucleotides, proteins, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells expressing the
CC polypeptide; methods for stimulating an immune response or treating
CC cancer in a patient by administering the composition and diagnostic kits
CC comprising at least one of the oligonucleotide of, or an antibody and a
CC detection reagent consisting of a reporter group. The polypeptides and
CC polynucleotides are useful as vaccines for the treatment or prevention of
CC lung cancer, and for diagnosis and monitoring of such cancer. The
CC polynucleotide, polypeptide and antigen presenting cells can be
CC used to stimulate or expand T cells specific for a tumorous protein.
CC The polynucleotides may be used as probes or primers for nucleic acid
CC hybridisation, and in the preparation of ribozyme molecules for
CC inhibiting expression of tumour polypeptides and proteins in tumour
CC cells. The present sequence is one of the 183 lung cancer associated
CC polynucleotides.
XX
SQ Sequence 2247 BP; 591 A; 606 C; 612 G; 438 T; 0 other;
Query Match 76.9%; Score 318.4; DB 24; Length 2247;
Best Local Similarity 88.6%; Pred. No. 2.2e-99;
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;
QY 3 AACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62
Db 282 AACTTACAGACAGCTGCAAAATTA-CGATTTGTACAAAAAGCATGCAACCTTCATCTTG 340
QY 63 TTGATATCTGGATATGATGTCGAGCTTCCGAGACATGGCCTTAACAGCTGGACCA 122
Db 341 TTGATATCTGGATATGATGTCGAGCTTCCGAGACATGGCCTTAACAGCTGGACCA 400
QY 123 GCACGGAGATCAGGGCTGTNCCGCTGGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182
401 CCACCGAGATCA-GTGTGTCGCGCTCGAAGCTGTCATCTCCTCCATCTACTATCAGTTG 459
QY 183 AACAGCGCTTCTCTTACTACACAGATACGCGTGGAGCAGTCCATCAGTCTCCATCTC 242
Db 460 AACAAAGCGCTTCTCTTACTACCAAAATAGTGTGGAACAATCTATCAGCTCCTCCTC 519
QY 243 AATTTCATGGTCCGCGCTTACACAGTACGCGCGGAGGCGATGTCGCTGTTTCAGNT 302
Db 520 AACTTTATGATGTCGTCATGATGATGATGATGATGATGATGATGATGATGATGATG 579
QY 303 AAAGCTATGTTAGCAACCATGTGTGTGTGGAATAATGCTGGACAAATGAGATACATTTTC 362
Db 580 AAAGCTATGTTAGCAACCATGTGTGTGTGGAATAATGCTGGACAAATGAGATGATGTTTC 639
QY 363 TCCAGATGTCAGATTCCTAATGGCTTAATGATGTTNGGAAAGCTTCACCAAGT 414
Db 640 TCCAGATGTCAGATTCCTAATGGCTTAATGATGATGATGATGATGATGATGATGATG 691
RESULT 2
ID AAS73627
XX AAS73627 standard; cDNA; 2518 BP.

AC AAS73627;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #9431.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
DR P-PSDB; ABG09440.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 9431; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2518 BP; 717 A; 627 C; 588 G; 586 T; 0 other;
Query Match 44.2%; Score 183; DB 23; Length 2518;
Best Local Similarity 68.1%; Pred. No. 2e-52;
Matches 280; Conservative 0; Mismatches 129; Indels 2; Gaps 2;
QY 4 ACGTACAGACAGCTTGCAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTGT 63
Db 164 ACCTACAGACAGCATGCAAGCT-TAGGTTTGTTCAGAGAAGTAATGCAATTTTCACCTGT 222
QY 64 TCATATCTGGATATGATGTCGAGCTTCCGAGACATGGCCTTAACAGCTGACACACAG 123
Db 223 GGACATATGGATATGTCAGAGCAATTCGGGAATAATGCTCTGAAACCTGGACCCAA 282
QY 124 CACGAGATCAGCGCTGTNCCGCTGGAGACCGTCATCTCGTCCATCTACTATCAGTTGA 183
Db 283 CACTGAACCTCA-ACGTGTCCCGCTTAGAGGCTGTCTCTCCACTATTTTACCAGCTCA 341

XX PS Claim 1; Page 115; 189pp; English.

XX CC The invention relates to an isolated polynucleotide comprising a sequence

CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),

CC or their fragments, homologues, variants or complements and their encoded

CC polypeptides. Also included are an expression vector comprising the

CC polynucleotide operably linked to an expression control sequence; a host

CC cell transformed or transfected with an expression vector of; an isolated

CC antibody, or its antigen-binding fragment that specifically binds to the

CC polypeptide; a method for detecting the presence of a cancer in a

CC patient; a fusion protein comprising at least the polypeptide; an

CC oligonucleotide that hybridises to the polynucleotide under moderately

CC stringent conditions; a method for stimulating and/or expanding T cells

CC specific for a tumour protein; an isolated T cell population comprising T

CC cells prepared from the method of above; a composition comprising a first

CC component consisting of carriers and immunostimulants, and a second

CC component selected from the polynucleotides, proteins, antibodies, fusion

CC proteins, T cell populations and antigen presenting cells expressing the

CC polypeptide; methods for stimulating an immune response or treating

CC cancer in a patient by administering the composition and diagnostic kits

CC comprising at least one of the oligonucleotide of, or an antibody and a

CC detection reagent consisting of a reporter group. The polypeptides and

CC polynucleotides are useful as vaccines for the treatment or prevention of

CC lung cancer, and for diagnosis and monitoring of such cancer. The

CC polynucleotide, polypeptide and antigen presenting cells can be

CC used to stimulate or expand T cells specific for a tumorous protein.

CC The polynucleotides may be used as probes or primers for nucleic acid

CC hybridisation, and in the preparation of ribozyme molecules for

CC inhibiting expression of tumour polypeptides and proteins in tumour

CC cells. The present sequence is one of the 183 lung cancer associated

CC polynucleotides.

XX SQ Sequence 512 BP; 127 A; 135 C; 141 G; 109 T; 0 other;

Query Match 43.7%; Score 181; DB 24; Length 512;

Best Local Similarity 87.0%; Pred. No. 4e-52;

Matches 221; Conservative 0; Mismatches 31; Indels 2; Gaps 2;

QY 3 AACGTACAGAACAGCTTGCAGTTCACGATTTGTACAGAACGCGATGCAACCTTCATCTTG 62

DB 260 AACTTACAGAACAGCTGCAATTA-CGATTTGTACAAAAGCATGCAACCTTCATCTTG 318

QY 63 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAAGTGGCTTAACAGCTGGACCACA 122

DB 319 TTGATATCTGGAACATGATGTAAGCCTTCCGAGACAAGTGGCTTAACAGCTGGACCACA 378

QY 123 GCACGAGATCAGGGGTGTTCGCTGGAGACCGCTCATCTCGTCATCTACTATCATCATGTTG 182

379 CCACCGAGATCA-GTGTGTCCCGCTCGAAACTGTCACTCTCCATCTACTATCATGTTG 437

QY 183 ACAAGCGCTTCTTCTACTCACAGATCAGCGTGGAGAGATCATCATGCTCTCTACTCTC 242

DB 438 ACAAGCGCTTCTTCTACTCACCAAAATAGTGTGGAACAATCTATCATGCGCTCTCTCTC 497

QY 243 AATTTATGCTGCGC 256

DB 498 AACTTTATGATTGC 511

RESULT 5

ABL14199

ID ABL14199 standard; cDNA; 3045 BP.

XX ABL14199;

AC ABL14199;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37079.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX

OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-658860/75.

DR P-PSDB; ABB70096.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Claim 1; SEQ ID NO 37079; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3045 BP; 721 A; 882 C; 846 G; 596 T; 0 other;

Query Match 36.4%; Score 150.8; DB 23; Length 3045;

Best Local Similarity 63.2%; Pred. No. 3.3e-41;

Matches 244; Conservative 0; Mismatches 141; Indels 1; Gaps 1;

QY 29 CGATTTGTACAGAACGCGATGCAACCTTCATCTTGTGTATATCTGATATGATCGAAGCT 88

DB 602 CGCTATATACAGAGTCCACAATCTGCACCTTGGTGACATATGGAATGTGATCGAGGCC 661

QY 89 TTCCGAGACAATGGCCTTTAACACGCTGGACCACACGACGAGATCAGCGGTGTCGGCT 148

DB 662 TTTCGCGAGATGGCCTGAACACCCCTGGAGCGCGACGAGGTGA-GCGTGGCCAGGCT 720

QY 149 GGAGACCGTCACTCGTCCATCTACTATCATCTTGNACAAAGCGCTTCCTTCTACTCACA 208

DB 721 GGAACCCCTGCTCTCTCTCTATCATTAACCTCAACAAGCGCTGCGCCACCGCTCACA 780

QY 209 GATCAGCGTGGAGAGTCCATCAGTCTCTACTCAATTTTCATGTCGCGCGCTACGACAG 268

DB 781 GGTGCTGTGACTCGAAGCGGGTCTGCTGCTCACTGCTGCTGCGCGGTACACAG 840

QY 269 TCAGGGCCGAGCGANGTTGACCGGTGTTTTCAGNTAAAGCTATGTTAGCAACCATGTTGG 328

DB 841 TGATAACTCGGGCAAGATACCGGTTTTCTCCATCAAAGTGGCCCTGGCCACTATGTGCTC 900

QY 329 TGGAAAATGCTGGACAAATTCAGATACATATTTCTCCAGATGTCAGATTCATATGGCTT 388

DB 901 TCGCAAGCTGGTGACAAACTAAGATACATATTTCTCGAGATCTCGGATGGCGTGGACA 960

QY 389 AATGATGTTNGGAAAGCTTGACCACT 414

DB 961 GCTGCTGCCCTGGAGAGCTGGCGGAGT 986

RESULT 6

ABL14198/c

ID	ABL14198	standard; cDNA; 15987 BP.
XX	AC	ABL14198;
XX	AC	ABL14198;
XX	AC	ABL14198;
DT	26-MAR-2002	(first entry)
XX	XX	
DE	Drosophila melanogaster	expressed polynucleotide SEQ ID NO 37076.
XX	XX	
KW	Drosophila; developmental biology; cell signalling; insecticide;	
KW	pharmaceutical; gene; ss.	
XX	XX	
OS	Drosophila melanogaster.	
XX	XX	
PN	WO200171042-A2.	
XX	XX	
PD	27-SEP-2001.	
XX	XX	
PF	23-MAR-2001; 2001WO-US09231.	
XX	XX	
P	23-MAR-2000; 2000US-191637P.	
Ph	11-JUL-2000; 2000US-0614150.	
XX	XX	
PA	(PEKE) PE CORP NY.	
XX	XX	
PI	Venter JC, Adams M, Li PWD, Myers EW;	
XX	XX	
DR	WPI; 2001-656860/75.	
DR	P-PSDB; ABB70095.	
XX	XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more	
PT	genes from Drosophila and for elucidating cell signalling and cell-cell	
PT	interactions -	
XX	XX	
PS	Claim 1; SEQ ID NO 37076; 21pp + Sequence Listing; English.	
XX	XX	
CC	The invention relates to an isolated nucleic acid detection reagent	
CC	capable of detecting 1000 or more genes from Drosophila. The invention is	
CC	useful in developmental biology and in elucidating cell signalling and	
CC	cell-cell interactions in higher eukaryotes for the development of	
CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA	
CC	sequences (ABL01840-ABL16175) and the encoded proteins	
CC	(ABB57737-ABB72072).	
CC	The sequence data for this patent did not form part of the printed	
CC	specification, but was obtained in electronic format directly from WIPO	
CC	at ftp.wipo.int/pub/published_pct_sequences.	
XX	XX	
SC	Sequence 15987 BP; 4295 A; 3570 C; 3827 G; 4295 T; 0 other;	
XX	XX	
XX	Very Match	22.8%; Score 94.2; DB 23;
XX	Best Local Similarity	64.6%; Pred. No. 3.6e-21;
XX	Matches 155; Conservative	0; Mismatches 84; Indels 1; Gaps 1;
QY	29 CGATTGTTACAGACGGATGCAACCTTCATCTGTTGATATCTGGAATATGATCGAAGCT	88
Db	4518 CGCTATATACAGAAGTCCACAAATGCGACTTGGTGGACATATGGAATGTGATCGAGGCC	4459
QY	89 TTCCGAGACATGCGCTTAAACAGCTGGACACACGAGACGAGAGTCAGGCGTGTNCCGCT	148
Db	4458 TTTCGCGAAGATGGCTGAACACCTGGAGCGCGAGAGCGAGGTGA -GGGTGGCCAGGCT	4400
QY	149 GGAGACGGTCAATCTCGTCCATCTACTATCAAGTGAACAGCGCCCTTCCTTCTACTCA	208
Db	4399 GGAACCCCTGGTCTCTCCCTCTATCATACCTCAACAAGCCCTGCCACCGCTCAGCA	4340
QY	209 GATCAGGCTGGAGCAGTCCATCAGTCTCCTACTCAATTTCAATGTCGGCGGCTACGACAG	268
Db	4339 GGTGCTGTGACTCGAAGGCGGGTCTGCTGCTCAACTGCTGCTGGCGGCTACACAAG	4280
XX	XX	
RESULT 7		
AAQ60778		
ID	AAQ60778	standard; DNA; 472 BP.


```

AC ABL87347;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:10325.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US17756.
XX
PR 26-MAY-2000; 2000US-207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
WPI; 2002-122075/16.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (III) of an ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour cDNA library using well known
CC techniques.
XX
SQ Sequence 689 BP; 201 A; 156 C; 161 G; 162 T; 9 other;
Query Match 18.9%; Score 78.4; DB 24; Length 689;
Best Local Similarity 83.3%; Pred. No. 1.8e-16;
Matches 85; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 313 TAGCAACCATGTGTGTGGAAAATGCTGGCAAAATGAGATACATTTTCTCCAGATGT 372
||||| ||||||||| ||||||||| ||||||||| ||| |||||||||
Db 688 TAGCACCATGTGTGTGGGANAATGCTGGCAATTCAGATATGTTTCTCCAGATGT 629
||||| ||||||||| ||||||||| ||||||||| ||| |||||||||

QY 373 CAGATTCCAATGGCTTAATGATGTGTTNGGAAGCTTGACCAGT 414
||||| ||||||||| ||||||||| ||||||||| ||| |||||||||
Db 628 CAGATTCCAATGGCTTAATGATATTTAGCAAGTGTGACCAGT 587
||||| ||||||||| ||||||||| ||||||||| ||| |||||||||

RESULT 9
AAC01269
ID AAC01269 standard; cDNA; 458 BP.
XX
AC AAC01269;

```

```

XX 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 1267.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
WPI; 2000-500381/45.
XX
DR P-PSDB; AAG01263.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 458 BP; 145 A; 77 C; 114 G; 121 T; 1 other;
Query Match 17.2%; Score 71.2; DB 21; Length 458;
Best Local Similarity 69.7%; Pred. No. 4.4e-14;
Matches 122; Conservative 1; Mismatches 50; Indels 2; Gaps 2;

QY 4 ACGTACAGAACAGCTTGCAAGTTACCGATTGTACAGAGCGATGCAACCTTCATCTTGT 63
||| ||||||||| ||||||| ||| ||||||||| ||| |||||||||
Db 286 ACCTACAGAACAGCATGCAAGCT-TAGGTTTGTTCAGAGAAATGCAATTTGCACCTGT 344
||||| ||||||||| ||||||||| ||||||||| ||| |||||||||

QY 64 TGATATCTGGAATATGATGGAAGCTTCCGAGACATGCGCTTAACACGCTGGACCACAG 123
||| ||||||||| ||||||||| ||||||||| ||| |||||||||
Db 345 GGACATATGGAATGTGATAGAGCATTTGGCGGAAATGCTCTGAGCAACCTGACCCAAA 404
||||| ||||||||| ||||||||| ||||||||| ||| |||||||||

QY 124 CAGCGAGATCAGCGGTGTNCCCGCTGGAGACCGTATCTCGTCCATCTACTATCA 178
||| ||||||||| ||||||||| ||||||||| ||| |||||||||
Db 405 CACTGAACTCA-ACGTTGCCGCTTAGAGCTGTGCTCTCCACTATTTTACC 458
||||| ||||||||| ||||||||| ||||||||| ||| |||||||||

RESULT 10
AAS31142
ID AAS31142 standard; cDNA; 5106 BP.
XX
AC AAS31142;
XX
DT 04-DEC-2001 (first entry)
XX

```



```
Query Match          9.2%; Score 38.2; DB 24; Length 13815;
Best Local Similarity 53.0%; Pred. No. 0.085;
Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 240 CTCATTTTCATGTCGCGCTAGACAGTGAGGCGCGAGCGANGTTGACCGGTGTTTCA 299
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9734 CTCACCTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCCGTGTCT 9793
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 300 GNTAAAGCTATGTTAGCAACCATGTCGTGGTGGAAATGCTGGCAAAATTCAGATACATT 359
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9794 TTTAAAGCTGGCATCATTTCTCTGTGTAAAGCACACTTGGAGACAGTAGTACATACCTT 9853
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 360 TTCTCCAGATGTCAGATTCATTCATGGCTT 388
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9854 TTCAAGCAAGTGGCAAGTTCAACTGGCTT 9882
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
AAT2558/C
I' AAT27558 standard; cDNA; 19307 BP.
XX
AC AAT27558;
XX
DT 07-AUG-1996 (first entry)
XX
DE Shuttle vector pAdDel.CMVDys.
XX
KW Adenovirus type 5; Ad5; vector; gene therapy; gene transfer;
KW helper virus; dystrophin; muscular dystrophy; ds; cyclic.
XX
OS Synthetic.
XX
XX WO9613597-A2.
XX
XX 09-MAY-1996.
XX
XX 27-OCT-1995; 95WO-US14017.
XX
XX 28-OCT-1994; 94US-0331381.
XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX
XX Chen S, Fisher KJ, Weitzman M, Wilson JM;
XX WPI; 1996-251463/25.
XX
XX Recombinant adenovirus produced from shuttle vector and helper virus
XX - has crippled packaging function, useful for delivering
XX transgene(s) to target cells
XX
XX Example 9; Fig 12A-12P; 149pp; English.
XX
XX Shuttle vector pAdDel.CMVDys (AAT27558) comprises the adenovirus
XX cis elements needed for replication and virion encapsidation but
XX is deleted of all viral genes. It carries a mouse dystrophin
XX minigene under control of the CMV promoter. It is used with a helper
XX virus that supplies the sequences needed for a productive viral
XX infection but which has disabled packaging function. Recombinant
XX adenovirus is produced that is characterized by high titer transgene
XX delivery to host cells and the ability to stably integrate the
XX transgene into the host cell chromosome. Such a vector can be used
XX to transfer the dystrophin gene for use in muscular dystrophy gene
XX therapy.
XX
XX Sequence 19307 BP; 4816 A; 4393 C; 4056 G; 6042 T; 0 other;

Query Match          9.2%; Score 38.2; DB 17; Length 19307;
Best Local Similarity 53.0%; Pred. No. 0.1;
Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 240 CTCATTTTCATGTCGCGCTAGACAGTGAGGCGCGAGCGANGTTGACCGGTGTTTCA 299
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 4951 CTCACCTGGCTTCTCAATGTTTATGATACGGGACGAACAGGAGGATCCCGTGTCT 4892
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 300 GNTAAAGCTATGTTAGCAACCATGTCGTGGTGGAAATGCTGGCAAAATTCAGATACATT 359
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9794 TTTAAAGCTGGCATCATTTCTCTGTGTAAAGCACACTTGGAGACAGTAGTACATACCTT 9853
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 360 TTCTCCAGATGTCAGATTCATTCATGGCTT 388
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9854 TTCAAGCAAGTGGCAAGTTCAACTGGCTT 9882
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
ABK81993
ID ABK81993 standard; DNA; 887 BP.
XX
AC ABK81993;
XX
DT 13-AUG-2002 (first entry)
XX
DE Human dystrophin C-terminal polynucleotide.
XX
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;
KW Duchenne's muscular dystrophy; DMD; dystrophin; human; ds.
XX
XX Homo sapiens.
XX
XX WO200229056-A2.
XX
XX 11-APR-2002.
XX
XX 04-OCT-2001; 2001WO-US31126.
XX
XX 06-OCT-2000; 2000US-238848P.
XX
XX (UNMI ) UNIV MICHIGAN.
XX
XX Chamberlain JS, Harper SQ;
XX WPI; 2002-435334/46.
XX
XX A composition for preparing therapeutic drugs, has a mini-dystrophin
XX peptide comprising a specific number of spectrin-like repeat domains,
XX or a nucleic acid sequence encoding the mini-dystrophin peptide
XX
XX Disclosure; Fig 9; 145pp; English.
XX
XX The invention describes a composition comprising a mini-dystrophin
XX peptide comprising a spectrin-like repeat domain, where the domain
XX comprises n spectrin-like repeats, and contains no more than n
XX spectrin-like repeats, where n is an even number between 4-24, or a
XX nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin
XX peptide or the polynucleotide encoding it is useful as a medicament,
XX for preparing a drug for therapeutic application and in the preparation
XX of a composition for treatment of muscle disease, e.g. Duchenne's
XX muscular dystrophy (DMD). This sequence represents a human dystrophin
XX polynucleotide sequence used in the creation of the mini-dystrophin
XX peptides of the invention.
XX
XX Sequence 887 BP; 244 A; 207 C; 213 G; 223 T; 0 other;

Query Match          9.0%; Score 37.2; DB 24; Length 887;
Best Local Similarity 50.0%; Pred. No. 0.039;
Matches 90; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 209 GATCAGCTGGAGCAGTCCATCAGTCTCTCTACTCTCAATTCATGTCGGCGCTAGACAG 268
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 168 GGTCAACGTCCCTCTCTCTGCTGGATATGTCGTAACCTGGCTGCTCAATGTTTATGATAC 227
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 269 TGAGGCGCGAGCAGTGGACCGTGTGTTTTCAGTAAAGCTATGTTAGCAACCATGTCGG 328
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 228 GGGACGAACAGGAGGATCCGCTGCTCTCTTTTAAACTGGCATCATTTCCCTGTGTAA 287
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 329 TGGAAAAATGCTGGACAAATTCAGATACATTTCTCCAGATGTCAGATTTCCAATGGCTT 388
    ||||| || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Db . 288 AGCACATTTTGAAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATT 347

Search completed: December 1, 2002, 16:24:34
Job time : 290 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
Run on: December 1, 2002, 16:20:13 ; Search time 53 seconds
(without alignments)
3008.260 Million cell updates/sec

Title: US-09-680-959-125
Perfect score: 414
Sequence: 1 ctaacgtacagaacagcttg.....gttnggaagcttgaccagt 414

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
341543 seqs, 192557720 residues

To. number of hits satisfying chosen parameters: 683086
Minimum DB seq length: 0
Maximum DB seq length: 0

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications, NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_NEW_PUB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
15: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

result No.	Score	Query Match	Length	ID	Description
1	318.4	76.9	2247	10	US-09-960-253-157
2	181	43.7	512	10	US-09-960-253-157
3	78.4	18.9	689	10	US-09-867-701-10325
4	37.2	9.0	13957	10	US-09-782-378A-22
5	37.2	9.0	13957	10	US-09-880-107-2284
6	32.2	7.8	1173	9	US-09-938-842A-1992
7	32	7.7	1040	10	US-09-825-345-1
8	30.8	7.4	80959	9	US-09-815-242-7522
9	30	7.2	1098	10	US-09-764-868-66
10	29.6	7.1	1927	9	US-09-198-559-2
11	29.6	7.1	2734	9	US-09-764-868-66
12	29.6	7.1	2964	10	US-09-198-559-2
13	29.2	7.1	307	10	US-09-294-093B-1641
14	29.2	7.1	348	10	US-09-864-761-21546
15	29.2	7.1	442	10	US-09-864-761-21546
16	28.8	7.0	2010	10	US-09-887-576-301
17	28.6	6.9	357	10	US-09-878-574-2805
18	28.6	6.9	402	10	US-09-777-564-462
19	28.6	6.9	408	10	US-09-878-574-4089

ALIGNMENTS
RESULT 1
US-09-960-253-157
; Sequence 157, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: 210121.556
; CURRENT APPLICATION NUMBER: US/09/960,253
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 157
; LENGTH: 2247
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-253-157

QY	3	ACGTACAGAACAGCTTGCAGCTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG	76.9%;	Score 318.4;	DB 10;	Length 2247;
Db	282	AACTTACAGAACAGCTTGCAGCTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG	88.6%;	Pred. No. 4.2e-99;	Mismatches 0;	Indels 2;
QY	63	TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGGCCTTACACGCTTGACACCA	122			
Db	341	TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGGCCTTACACGCTTGACACCA	122			
QY	123	GCACCGAGATCAGGCGTTCGAGCTTTCGAGACAAATGGCCTTACACGCTTGACACCA	400			
Db	401	CCACCGAGATCAGGCGTTCGAGCTTTCGAGACAAATGGCCTTACACGCTTGACACCA	400			
QY	183	AAACAGCGCTTCTCTTACTACACGATCAGGCTGAGAGCTTCTCTCTCTCTCTCTCTCT	182			
Db	460	AAACAGCGCTTCTCTTACTACACGATCAGGCTGAGAGCTTCTCTCTCTCTCTCTCTCT	459			
QY	243	AATTTATGGTGGCGCTACGACAGTGGCGCGGAGGCAAGTGGACGGTGTTCAGNT	302			

ue Dec 3 12:27:59 2002

520 AACTTTATGATTGCTGCATATGACAGTGGGCCGAGCAAGTTGCGGTATTTTCAGTT 579
 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAATAATGCTGGACAAATTGAGATACATTTTC 362
 580 AAAGCTATGTTAGCAACCATGTGTGGTGGAATAATGCTGGACAAATTGAGATATGTTTC 639
 363 TCCGAGATGTCAGATTCCCAATGCTTAAATGATCTTTNGGAAAGCTTGCCACAGT 414
 640 TCCGAGATGTCAGATTCCCAATGCTTAAATGATCTTTAGCAAGTTTGACCAGT 691

```

;RESULT 2
;Sequence 28, Application US/09960253
Patent No. US20020123619A1
GENERAL INFORMATION:
APPLICANT: Benson, Darin R.
APPLICANT: Monamath, Raodoh
APPLICANT: Lodes, Michael J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
-TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
TITLE OF INVENTION: A 556
FILE REFERENCE: 210121, 556
CURRENT APPLICATION NUMBER: US/09/960,253
CURRENT FILING DATE: 2001-09-20
NUMBER OF SEQ ID NOS: 187
SOFTWARE: FastSeq for Windows version 4.0
SEQ ID NO 28
LENGTH: 512
TYPE: DNA
ORGANISM: Homo sapiens
US-09-960-253-28

```

[illegible]

RESULT 3
US-09-867-701-10325/C
; Sequence 10325, Application US/09867701
; Patent No. US2002013227A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; OF OVARIAN CANCER
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; TITLE OF INVENTION: 210121.497
; FILE REFERENCE: US/09/867,701
; CURRENT APPLICATION NUMBER: 2001-05-29
; CURRENT FILING DATE: 19912
; NUMBER OF SEQ ID NOS: 4.0
; NUMBER OF SEQ ID NOS: 4.0
; NUMBER OF SEQ ID NOS: 4.0

```
; SEQ ID NO 10325
; LENGTH: 689
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(689)
; OTHER INFORMATION: n = A,T,C or G
; US-09-867-701-10325
Query Match      18.9%; Score 78.4; DB 10; Length 689;
Best Local Similarity 83.3%; Pred. No. 5.6e-17; Indels 0; Gaps 0;
Matches 85; Conservative
QY   313 TAGCAACCATGTGCTGGGAAAATGCTGCACAAATTGAGATACATATTTTCCCAGATGT 372
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    688 TAGCACCATGTGTGGTGGANNAATGCTGGCAAAATTGAGANTATGTTNTCTCCCAGATGT 629
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY   373 CAGATTTCCAATGGCTTAATGATGTTNGAAAGCTTGACCAGT 414
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY   628 CAGATTTCCAATGGCTTAATGATGTTTAGCAAGTNTGACCAGT 587
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 4
US-09-782-378A-22
Sequence 22, Application US/09782378A
Patent No. US20020102731A1
GENERAL INFORMATION:
APPLICANT: Hearing, Patrick
APPLICANT: Bahou, Wadie
APPLICANT: Gandelsona, Zliv
APPLICANT: Gnatsenko, Dmitri
APPLICANT: Adenoviral Vectors
TITLE OF INVENTION: Adenov04970
FILE REFERENCE: STONYB-04970
CURRENT APPLICATION NUMBER: US/09782, 378A
CURRENT FILING DATE: 2001-02-12
CURRENT APPLICATION NUMBER: 60/237,747
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 27
SOFTWARE: Patentin version 3.0
SEQ ID NO 22
LENGTH: 13957
TYPE: DNA
ORGANISM: Homo sapiens
US-09-782-378A-22

```

	Query Match	9.0%;	Score 37.2;	Db 10,		Gaps
	Best Local Similarity	50.0%;	Pred. No. 0.05;			
Matches	90; Conservative	0; Mismatches	90; Indels			
209	GATCAGCGTGGAGACGTC	CATCAGTCTCTACTCAATTC	TCAATTC	TGCTGC	GTCCGCGCTACGACAG	268
QY						
9712	GGTCAACGTCCTCTCTCG	TGGATATGTC	TCTGAAC	TGSGCTG	CTCAATGTTTATGATAC	9771
Ddb						
269	TGAGGCGCGAGCGANGIT	GACCGTGTTTTTCAGNTPA	AAGCTATGTTAGCA	ACCATTGTGTG	GG	328
QY						
9772	GGACGAACAGGAGGATCC	GTGCCGTCTCTTTAA	AACTGGCA	TCAATTC	CCCTGTGTA	9831
Db						
329	TGGAATAATTC	TGGACAATTCAGATAC	ATTTC	TCCCAGATGTCAGAT	TCCAATGCGTT	388
QY						
9832	AGCATTTGGGAAGACAG	ATACAGATACCTTTT	TTCAGCAAGTGGCA	AGTTCAACAGGAT		9891
Db						

```

RESULT 5
US-09-880-107-2284
US-09-880-107-2284, Application US/09880107
; Sequence 2284, Application US/09880107
; Patent NO. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; APPLICANT: Gene Expression Profiles in Liver Cancer
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

```

	Query Match	7.8%;	Score 32.2;	DB 9;	Length 1173;
	Best Local Similarity	47.9%;	Pred. No. 0.57;		
	Matches	91;	Conservative 0;	Mismatches 99;	Indels 0; Caps 0;
QY	15	AGCTTGCAAGTTACAGAGCGGTGCAACCTTCATCTTGGTATATCGGA	74		
Db	206	ACCTTGTAAGCTCCGGTGAAGCCAGGAGAGATCGAAGCTCTTGGAGACAGTGAAGA	147		
QY	75	ATATGATCGAAGCTTCCGAGACAAATGCCCTTAACAGCTGGACCAAGCAGGAGATCA	134		
Db	146	ATAGGATAGTATAATTTAGAGAGCGGTGCCATGATGACCAATACCAATCATCGGGAAG	87		

```

RESULT 8
US-09-858-546-3
; Sequence 3, Application US/09858546
; Patent No. US2002017295A1
; GENERAL INFORMATION:
; APPLICANT: SHAO, Wei et al
; TITLE OF INVENTION: ISOLATED HUMAN SECRET
; TITLE OF INVENTION: NUCLEIC ACID MOLECUL
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001232
; CURRENT APPLICATION NUMBER: US/09/858,546
; CURRENT FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 80959
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(80959)
; OTHER INFORMATION: n = A,T,C or G

```


Patent No. US20020025930A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: RAS-LIKE PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,559
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/846,790
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0388 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2964 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNUT16
CLONE: 2791521
US-09-198-559-2

Query Match 7.1%; Score 29.6; DB 10; Length 2964;
Best Local Similarity 53.0%; Pred. No. 7.9;
Matches 62; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 58 TCTTGTGATATCTGATATGATGAGCTTTCGAGACAATGGCCTTAACAGCTGGA 117
DB 863 TCTCTTTTCATATACAGCAACGGGCAAGGTTTGTGGTCTATCAGGAGATCAATCTGGG 804
QY 118 CCACAGCAGGAGATCAGCGGTGCCCTCGAGACCTCATCTCGTCCATCTACT 174
DB 803 TCCCTATGAGGACATAGGACGCTGAGGATCGATGCTTGTAGCTCGGGGACCCATT 747

RESULT 13
US-09-294-093B-1641
Sequence 1641, Application US/09294093B
Patent No. US20010051335A1
GENERAL INFORMATION:
APPLICANT: Lalgudi, Raghunath, V.
APPLICANT: Ito, Laura, Y.
APPLICANT: Sherman, Bradley, K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
FILE REFERENCE: PL-0009 US
CURRENT APPLICATION NUMBER: US/09/294,093B
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: 60/082,567
PRIOR FILING DATE: April 21, 1998
NUMBER OF SEQ ID NOS: 6207

SOFTWARE: PERL Program
SEQ ID NO 1641
LENGTH: 307
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20010051335A1 700344740H1
NAME/KEY: unsure
LOCATION: 12-13, 30-31, 41, 43-44, 70, 79, 81, 105-106, 140, 153, 178, 192, 196,
OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-1641
Query Match 7.1%; Score 29.2; DB 10; Length 307;
Best Local Similarity 51.8%; Pred. No. 2.7;
Matches 58; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
QY 139 TGTNCGCCTCGAGACCGTCATCTGTCATCTACTATCATCTAGTGAACAAGCGCTTCCTT 198
DB 146 TATATCAGTTGACGGGACAGAGATCCATCTCTATAAGCGCTNAANAATGCCCTCTG 205
QY 199 CTACTCACCAGATCAGCGGTGAGCAGTCCATCATCTCTCTACTCTCAATTCAT 250
DB 206 CGATTAAACAGTTCACCCAGGCTGGACCGCAACAGCTATNCAGCTGCTT 257

RESULT 14
US-09-864-761-21546/C
Sequence 21546, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 16:14:38 ; Search time 1993 Seconds
(without alignments)
3381.206 Million cell updates/sec

Title: US-09-680-959-125

Perfect score: 414

Sequence: 1 ctaactacagaaacagcttg.....gttnggaagcttgaccagt 414

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

1 . number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em.estba:*

2: em.esthum:*

3: em.estin:*

4: em.estmu:*

5: em.estov:*

6: em.estpl:*

7: em.estro:*

8: em.htc:*

9: gb.est1:*

10: gb.est2:*

11: gb.htc:*

12: gb.est3:*

13: gb.est4:*

14: gb.est5:*

15: em.estfun:*

16: em.estom:*

17: gb.gss:*

18: em.gss_hum:*

19: em.gss_inv:*

20: em.gss_pln:*

21: em.gss_vrt:*

22: em.gss_fun:*

23: em.gss_mam:*

24: em.gss_mus:*

25: em.gss_other:*

26: em.gss_pro:*

27: em.gss_rdo:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	382.4	92.4	505	10	BE650744 UI-M-BH3-
2	363	87.7	517	9	AI325959 mq63d07.x
3	348.2	84.1	958	12	BG865074 602784226
4	330.8	79.9	676	10	BB642867 BB642867
5	328	79.2	461	10	AW496140 up47e06.y
6	318.4	76.9	791	12	BG700652 602682269

7	318.4	76.9	880	14	BQ423333
8	318.4	76.9	932	14	BQ651659
9	318.4	76.9	967	9	AL524324
10	318	76.8	896	9	AL555818
11	317	76.6	845	13	BI854854
12	311.4	75.2	647	10	BB654888
13	307	74.2	590	12	BG078688
14	303	73.2	600	12	BG805399
15	299.4	72.3	830	12	BG469017
16	299	72.2	600	12	BG804143
17	299	72.2	622	10	BE500997
18	295	71.3	448	10	AW495863
19	293	70.8	673	10	BB640953
20	290.8	70.2	823	9	AL527400
21	290.4	70.1	571	10	BE552374
22	279.6	67.5	760	9	AL538124
23	277.2	67.0	915	12	BE779121
24	273.6	66.1	595	9	AJ454452
25	270.4	65.3	712	10	AW327909
26	261.6	63.2	954	12	BG115265
27	260.8	63.0	584	13	BI339696
28	255	61.6	654	13	BJ065983
29	245	59.2	926	9	AL530827
30	243	58.7	878	12	BG702733
31	242	58.5	419	12	BE947648
32	241	58.2	457	12	BE950528
33	240.2	58.0	422	10	AW491035
34	212	51.2	381	12	BF721737
35	210.4	50.8	541	9	AL636722
36	210	50.7	352	10	BE113690
37	194.6	47.0	390	10	AW379323
38	194	46.9	464	10	BE482496
39	194	46.9	574	12	BG689246
40	190.8	46.1	396	9	AI574752
41	190	45.9	1061	14	BN918263
42	186.2	45.0	918	14	BQ891601
43	186.2	45.0	1608	11	AK003829
44	186.2	45.0	3861	11	AK019477
45	184.2	44.5	701	13	BI546456

ALIGNMENTS

RESULT 1
BE650744/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

BE650744
UI-M-BH3-asr-h-05-0-UI-r1 NIH-BMAP_M_S4 Mus musculus cDNA clone
UI-M-BH3-asr-h-05-0-UI 5', mRNA sequence.
BE650744.1 GI:9976616
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 505)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mstet@mail.nih.gov
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose

880 14 BQ423333
932 14 BQ651659
967 9 AL524324
896 9 AL555818
845 13 BI854854
647 10 BB654888
590 12 BG078688
600 12 BG805399
830 12 BG469017
600 12 BG804143
622 10 BE500997
448 10 AW495863
673 10 BB640953
823 9 AL527400
571 10 BE552374
760 9 AL538124
915 12 BE779121
595 9 AJ454452
712 10 AW327909
954 12 BG115265
584 13 BI339696
654 13 BJ065983
926 9 AL530827
878 12 BG702733
419 12 BE947648
457 12 BE950528
422 10 AW491035
381 12 BF721737
541 9 AL636722
352 10 BE113690
390 10 AW379323
464 10 BE482496
574 12 BG689246
396 9 AI574752
1061 14 BN918263
918 14 BQ891601
1608 11 AK003829
3861 11 AK019477
701 13 BI546456

availability will be considered under appropriate and limited collaborative arrangements

```
FEATURES
  source
    Location/Qualifiers
      1..505
        /organism="Mus musculus"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UT-M-BH3-asr-h-05-0-UI"
        /clone_lib="NIH_BMAP_M_S4"
        /dev_stage="27-32 days"
        /lab_host="DH10B (Life Technologies)"
        /note="vector: pT73D-Pac (Pharmacia) with a modified
        polylinker: Site_1: Not I; Site_2: Eco RI; The
        NIH_BMAP_M_S4 library is a subtracted library of a series,
        ultimately derived from a mixture of individually tagged
        normalized libraries from ten regions of the mouse brain
        (cerebellum, brain stems, olfactory bulbs, hypothalamus,
        cortex, amygdala, basal ganglia, pineal gland, striatum,
        hippocampus) after a series of subtractions to reduce the
        representation of cDNAs from which ESTs had already been
        generated. The following serially subtracted libraries
        were generated in this process: NIH_BMAP_M_S4,
        NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S1,
        NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
        (NIH_BMAP_M_S4) was constructed as follows: PCR amplified
        cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
        NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
        was used as a driver in a hybridization with a pool of
        the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
        libraries in the form of single-stranded circles. The
        remaining single-stranded circles (subtracted library)
        was purified by hydroxyapatite column chromatography,
        converted to double-stranded circles and electroporated
        into DH10B bacteria (Life Technologies) to generate the
        NIH_BMAP_M_S4 library. This procedure has been previously
        described (Bonaldo, Lennon and Soares, Genome Research
        6:791-806, 1996)."
      134 a 113 c 128 g 130 t

BASE COUNT      134 a 113 c 128 g 130 t
ORIGIN
  Query Match      92.4%; Score 382.4; DB 10; Length 505;
  Best Local Similarity 98.3%; Pred. No. 1.8e-107;
  Matches 405; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Qy 3 AACGTACAGACAGCTTCCAGTAGTACCAGTTGTACAGAGCGATGCAACCTTCATCTTG 62
Db 498 AACGTACAGACAGCTTCCCAAGTTA-CGATTGTGTACAGAGCGATGCAACCTTCATCTTG 440

63 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAATGGCTTAACACGCTGGACCACA 122
Db 439 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAATGGCTTAACACGCTGGACCACA 380

Qy 123 GCACGGAGATCAGCGGTGTCGCGTGGAGACCGCTCATCTCGTCCATCTACTATCAGTTG 182
Db 379 GCACGGAGATCA-CGGTGTCCCGCTGGAGACCGCTCATCTCGTCCATCTACTATCAGTTG 321

Qy 183 AACAGGCCCTTCCTTCTTACTACACAGATCAGCTGGAGAGCTCCATCAGTCTCCTACTC 242
Db 320 AACAGGCCCTTCCTTCTTACTACACAGATCAGCTGGAGAGCTCCATCAGTCTCCTACTC 261

Qy 243 AATTTCATGGTCGCGCTACGACAGTACGGCGCGAGCGCANGTTGACCGTGTGTTTCAGNT 302
Db 260 AATTTCATGGTCGCGCTACGACAGTACGGCGCGAGCGCANGTTGACCGTGTGTTTCAGNT 201

Qy 303 AAAGCTATGTTAGCAACCATGTTGTGGTGGAAAAATGCTGGCAAAATGAGATACATTTTC 362
Db 200 AAAGCTATGTTAGCAACCATGTTGTGGTGGAAAAATGCTGGCAAAATGAGATACATTTTC 141

Qy 363 TCCAGATGTCAGATTCAGATGGCTTATGATGTTNGGAAGCTTGACCACT 414
Db 140 TCCAGATGTCAGATTCAGATGGCTTATGATGTTNGGAAGCTTGACCACT 89
```

RESULT 2

```
AI325959
LOCUS
  DEFINITION
    mg63d07.x1 Soares_thymus_2NDMT Mus musculus cDNA clone IMAGE:583405
    3' similar to TR:Q13197 Q13197 DYSTROBIN-DELTA. [1] ; mRNA
    sequence.
  ACCESSION
    AI325959
  VERSION
    AI325959.1 GI:4060388
  KEYWORDS
    EST.
  SOURCE
    house mouse.
  ORGANISM
    Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  REFERENCE
    1 (bases 1 to 517)
    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
    Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
    Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
    Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
    Waterston,R.
    The WashU-HMMI Mouse EST Project
    Unpublished (1996)
    Contact: Marra M/Mouse EST Project
    WashU-HMMI Mouse EST Project
    Washington University School of Medicine
    4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
    Tel: 314 286 1800
    Fax: 314 286 1810
    Email: mouseest@watson.wustl.edu
    This clone is available royalty-free through LLNL; contact the
    IMAGE Consortium (info@image.llnl.gov) for further information.
    MGI:358053
    This clone was previously sequenced on the 5' end only, this new
    data is from the 3' end
    Possible reversed clone: similarity on wrong strand
    High quality sequence stop: 463.

FEATURES
  source
    1..517
      /organism="Mus musculus"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="IMAGE:583405"
      /clone_lib="Soares_thymus_2NDMT"
      /sex="male"
      /tissue_type="Thymus"
      /dev_stage="4 weeks"
      /lab_host="DH10B"
      /note="vector: pT73D-Pac (Pharmacia) with a modified
      polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
      was primed with a Not I - oligo(dT) primer [5',
      TGTACCAATCTGAAGTGGAGCGCGCGCTTTTGTGTTTGTGTTTGTGTTTGTGTTT
      3']; double-stranded cDNA was ligated to Eco RI adaptors
      (Pharmacia), digested with Not I and cloned into the Not I
      and Eco RI sites of the modified pT73 vector. RNA
      provided by Dr. Bertrand Jordan. Library went through two
      rounds of normalization, and was constructed by Bento
      Soares and M. Fatima Bonaldo."
```

```
BASE COUNT      139 a 126 c 129 g 123 t
ORIGIN
  Query Match      87.7%; Score 363; DB 9; Length 517;
  Best Local Similarity 98.7%; Pred. No. 1.9e-101;
  Matches 385; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

Qy 3 AACGTACAGACAGCTTCCAGTTGACGATTTGTACAGAGCGATGCAACCTTCATCTTG 62
Db 130 AACGTACAGACAGCTTCCCAAGTTA-CGATTGTGTACAGAGCGATGCAACCTTCATCTTG 188

Qy 63 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAATGGCTTAACACGCTGGACCACA 122
Db 189 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAATGGCTTAAACACGCTGGACCACA 248

Qy 123 GCACGGAGATCAGCGGTGTCGCGCTGGAGACCGCTCATCTCGTCCATCTACTATCAGTTG 182
```

QY	123	GCACGGAGATCAGCGGTGTCNCCGCTCGGAGACCGTCACTCTCGTCATCTACTATCATGTTG	182
Db	411	GCACGGAGATCA - GCGTGTCGCCGCTCGGAGACCGTCACTCTCGTCATCTACTATCATGTTG	469
QY	183	AACAAGCGCCTTCCTTCTACTCACGATCACAGTGGAGCAGTCCATCAGTCTCTCTACTC	242
Db	470	AACAAGCGCCTTCCTTCTACTCACGATCACAGTGGAGCAGTCCATCAGTCTCTCTACTC	529
QY	243	AATTTCATGTCGCGCCTACGACAGTGGAGCGGCGAGGC - ANGTTGACCGTGTTTTTCAGN	301
Db	530	AATTTCATGTCGCGCCTACGACAGTGGAGCGGCGAGGC - ANGTTGACCGTGTTTTTCAGT	589
QY	302	TAAAGCTATGTTGAACACCACATGTGTGGTGGAAAAA -- TGCTTGGACAAATTCAGATACATT	359
Db	590	TAAAGCTATGTTGAACACCACATGTGTGGTGGAAAAAATGCTGTGACCAATTCAGATACATT	649
QY	360	TTCTCCAGATGTCAGATTCATCGGCTTAATGATGTTNGGAAAGCTTGACCACT	414
Db	650	TTCTCCAGATGTCAGATTCATCGGCTTAATGATGTTNGGAAAGCTTGACCACT	704
RESULT 4		676 bp mRNA linear EST 26-OCT-2001	
BB642867		BB642867 RIKEN full-length enriched, adult retina Mus musculus cDNA clone A930104E03 5', mRNA sequence.	
DEFINITION		BB642867 GI:16477625	
ACCESSION		BB642867.1	
VERSION		EST.	
KEYWORDS		house mouse	
SOURCE		Mus musculus	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 676)	
REFERENCE		Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.	
AUTHORS		RIKEN Mouse ESTs (Arakawa,T., et al. 2001)	
TITLE		Contact: Yoshihide Hayashizaki	
JOURNAL		Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute	
COMMENT		The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsr.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,R., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (10), 1757-1771 (2000) Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaoka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.	

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.9sc.riken.go.jp>) for further details.
e mouse tissues.

FEATURES
source

Location/Qualifiers
1. 676
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A930104E03"
/clone_lib="RIKEN full-length enriched, adult retina"
/tissue_type="retina"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: Sall; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5'GAGAGAGATCTCGAGTAATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. -Retina RNA was provided by Stefano Gustincich, Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA02115, USA, whose assistance we gratefully acknowledge."
150 a 180 c 200 g 144 t 2 others

BASE COUNT 150 a 180 c 200 g 144 t 2 others
ORIGINQuery Match 79.9%; Score 330.8; DB 10; Length 676;
Best Local Similarity 97.2%; Pred. No. 2.1e-91;
Matches 376; Conservative 0; Mismatches 7; Indels 4; Gaps 4;

QY 3 AACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62
|||||
Db 280 AACGTACAGACAGCTTGCAAGTTA-CGATTTGTACAGAGCGATGCAACCTTCATCTTG 338
QY 63 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAATGGCCTTAACACGCTGGACCACA 122
Db 339 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAATGGCCTTAACACGCTGGACCACA 398
123 GCACGGAGATCAGCGCGTGTNCCGCTGGAGACGCTCATCTCGTCCATCTACTATCAGTTG 182
Db 399 GCACGGAGATCA-GCGTGTCCGCTGGAGACGCTCATCTCGTCCATCTACTATCAGTTG 457
QY 183 AACAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTTC 242
Db 458 AACAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTTC 517
QY 243 AATTTCATGGTCCGCCCTACGACAGTGGAGCGGAGGCGGAGGAGTTGACCGTGTTCAGNT 302
Db 518 AATNTCATGGTCCGCCCTACGACAGTGGAGCGGAGGCGGAGGAGTTGACCGTGTTCAGNT 577
QY 303 AAAGCTATGTAGCAACCATGTGTGTGGAAAATGCTGGACAAATGTAGATACATTTTC 362
Db 578 AAAGCTATGTAGCAACCATGTGTGTGGAAAATGCTGGACAAATGTAGATACATTTTC 636
QY 363 TCCAGATGTCAGATTCGAATGGCTTA 389
Db 637 T-CCAGATGTCAGATTCGAATGGCTTA 662

RESULT 5
AW496140/c
LOCUS

461 bp mRNA linear EST 24-FEB-2000

DEFINITION

up47e06.y1 Soares_mouse_NMGB_bcell Mus musculus cDNA clone
IMAGE:2749282 5' similar to TR:070585 070585 DYSTROBREVIN, BETA ;,
mRNA sequence.

ACCESSION AW496140
VERSION AW496140.1 GI:7066421
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 461)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Other_ESTs: up47e06.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1035654

Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco.

FEATURES

source

1. 461
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:2749282"
/clone_lib="Soares_mouse_NMGB_bcell"
/lab_host="DH10B (phage-resistant)"
/note="Organ: germinal B-cell; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGGCCCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized; constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 107 a 117 c 109 g 128 t
ORIGINQuery Match 79.2%; Score 328; DB 10; Length 461;
Best Local Similarity 98.6%; Pred. No. 1.2e-90;
Matches 350; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 3 AACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62
|||||
Db 353 AACGTACAGACAGCTTGCAAGTTA-CGATTTGTACAGAGCGATGCAACCTTCATCTTG 295
QY 63 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAATGGCCTTAACACGCTGGACCACA 122
Db 294 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAATGGCCTTAACACGCTGGACCACA 235
QY 123 GCACGGAGATCAGCGCGTGTNCCGCTGGAGACGCTCATCTCGTCCATCTACTATCAGTTG 182
Db 234 GCACGGAGATCA-GCGTGTCCGCTGGAGACGCTCATCTCGTCCATCTACTATCAGTTG 176
QY 183 AACAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTTC 242
Db 175 AACAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTTC 116
QY 243 AATTTCATGGTCCGCCCTACGACAGTGGAGCGGAGGCGGAGGAGTTGACCGTGTTCAGNT 302
Db 115 AATTTCATGGTCCGCCCTACGACAGTGGAGCGGAGGCGGAGGAGTTGACCGTGTTCAGNT 56
QY 303 AAAGCTATGTAGCAACCATGTGTGTGGAAAATGCTGGACAAATGTAGATACATA 357
Db 55 AAAGCTATGTAGCAACCATGTGTGTGGAAAATGCTGGACAAATGTAGATACATA 1

RESULT 6

BC700652	791 bp	mrna	linear	EST 07-MAY-2001
LOCUS	CG2682269f1	NIH_MGC_95	Human sapiens	cdna IMAGE:4815064 5',
DEFINITION	mrna sequence.			
ACCESSION	BG700652			
VERSION	BG700652.1			
KEYWORDS	GI:13970209			
SOURCE	EST.			
ORGANISM	human.			
REFERENCE	Homo sapiens			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
JOURNAL	1 (bases 1 to 791)			
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ .			
	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgapbs-remail.nih.gov			
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.			
	CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki			
	Toshiyuki and Piero Carninci (RIKEN)			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLAM10711 row: i column: 17			
	High quality sequence stop: 761.			

```

FEATURES
source
1. .791
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4815064"
/clone_lib="NIH_MGC_95"
/tissue_type="hippocampus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI/XhoI (gtcagag
); Oligo-dT primed using primer 5'-tttttttttttttVN-3',
size-selected for average insert size 2.5 kb and
size-selected for ROR 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIMH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
130 c 215 g 181 t
205 a
BASE COUNT

```

	ry Match	76.9%;	Score 318.4;	DB 12;	Length 791;
	t Local Similarity	88.6%;	Pred. No. 1.6e-87;		
	atches 365;	Conservative 0;	Mismatches 45;	Indels 2;	Gaps 2;
QY	3	ACTACAGAACAGCTTGCAGT	TACCGATT	TGTACAGAGCGATGCAACCTT	CATCTTG 62
Db	275	AACCTTACAGAACAGCTGCAATTA	-CGATT	TGTACAAACAGTGCACCTT	CATCTTG 333
QY	63	TTGATATCTCGAATATATGATCGAAGCTT	TCGAGACAAATGGCTT	TACACAGCTGGACACACA	122
Db	334	TTGATATCTGGAACATGATTGAAGCCTT	CCGAGACAATGGCTT	TAATACACTGGACCAT	393
QY	123	GCACGGAGATCAGCGGTGNC	CGCTGGAGACCGTCATCTCGTCCATCT	TACTATCAGTTG	182
Db	394	CCACCGAGATCA-GTG	TGTCCCGCTCGAAACTGTCACTCTCCATCT	TACTATCAGTTG	452
QY	183	AACAAGCGCCTTCTTCT	TACTACACAGATCAGCTGGAGCATCCATCAGTCTCCTACTC	242	
Db	453	AACAAGCGCCTTCTTCT	TACTACAAATTAAGTGTGGAACAATCTATCAGCTCCTCCTC	512	
QY	243	AATTTATGTCGCCGCTTACGACAGT	GAGGGCGGAGGTCAGTGTGACCGTGT	TTTCAGNT 302	
Db	513	AACTTTATGATTCGTGCATATGACAGT	GAGGGCGGACAGTTGACGGTATTTTCAGTT	572	
QY	303	AAAGCTATGTTAGCACCATG	TGTGTGGAAAAATGCTGGACAAATTCAGATCATTTTC	362	

Db 573 AAAGCTATGTTAGCAACCATGCTGCTGCGAAAAAATGCTGGACAAATGAGATATGTTTC 632

Qy 363 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTNGGAAAGCTTCACCACT 414

Db 633 TCCAGATGTCAGATCCCAATGGCTTAATGATATTTAGCAAGTTTGACCACT 684

RESULT 7

BQ423333

LOCUS

DEFINITION BQ423333 880 bp mRNA linear EST 23-MAY-2000

AGENCOURT.7846069 NIH_MGC_72 Homo sapiens cdna clone IMAGE:6064183

5', mRNA sequence.

ACCESSION BQ423333

VERSION BQ423333.1 GI:21118648

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 880)

NIH-MGC <http://mgc.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC/DC/DTP

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution by: Agencourt Bioscience Corporation

found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM1338 Row: h column: 08

High quality sequence/posers

Location/Qualifiers

1. 880

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6064183"

/clone_lib="NIH_MGC_72"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pCMV-SPOPT6; Site:1: NotI; Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."

FEATURES

source

BASE COUNT	237 a	211 c	204 g	227 t	1 others
ORIGIN					
Query Match	76.9%	Score 318.4;	DB 14;	Length 880;	
Best Local Similarity	88.6%;	Pred. No. 1.7e-87;			
Matches 365;	Conservative 0;	Mismatches 45;	Indels 2;	Gaps 2;	
QY	3	AACCTACAGAACACTGCGAAGTTACCGATTTGACAGAACGGATGCAACCTTTCATCTTG	62		
Db	168	AACCTACAGAACGCCGTCGAAATTA-CGATTTGTACAAAACGATGCAACCTTTCATCTTG	226		
QY	63	TTGATATCTGGAATATGATCGAAGCTTTCGAGAGCAATGGCTTTAAACAGCTGGACACCA	122		
Db	227	TTGATATCTGGAACATGATTTGAAGCCTTCGAGAGCAATGGCTTTAATACATGGACCAT	286		
QY	123	GCAGGAGATCAGCGGTGTCNCCGCTCGAGACCGTCACTCGTCCATCTACTATCAGTTG	182		
Db	287	CCACCGAGATCA-GTGTGTCCCGCTCGAAACTGTCACTCTCTCCATCTACTATCAGTTG	345		
QY	183	AACAAGCGCCTTCCTTCTACTACAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC	242		
Db	346	AACAAGCGCCTTCCTTCTACTACCAAAATTAGTGTGGAACAATCATCAGCCTCCTCCTC	405		
QY	243	AAATTATGTCGCGCCTACGACAGTGAGGCGGAGCGANGTTGACCGTGTTCCTAGNT	302		
Db	406	AACCTTATGATGTCGATATGACAGTGAGGCGGAGCGAAGTTGACGGTATTTTCAGTT	465		

QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 362
 |||||
 Db 466 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATATGTTTC 525
 |||||
 QY 363 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTNGGAAAGCTTGACCAGT 414
 |||||
 Db 526 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTTAGCAAGTTTGACCAGT 577
 |||||

RESULT 8
 BQ651659 932 bp mRNA linear EST 15-JUL-2002
 LOCUS AGENCOURT_8342332 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6268278
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ651659
 VERSION BQ651659.1 GI:21775831
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 932)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CGAP (Stanford)
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Title Sequencing by: Agencourt Bioscience Corporation
 JOURNAL Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM2439 row: p column: 07
 High quality sequence stop: 733.
 FEATURES Location/Qualifiers
 1..932
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6268278"
 /clone_lib="NIH_MGC_100"
 /tissue_type="hepatocellular carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pOTB7; Site_1: xhoI; Site_2:
 EcoRI; cDNA made by oligo-dr priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCAGGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."
 BASE COUNT 225 a 233 c 252 g 221 t 1 others
 ORIGIN
 Query Match 76.9%; Score 318.4; DB 14; Length 932;
 Best Local Similarity 88.6%; Pred. No. 1.8e-87;
 Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;
 QY 3 AACGTACAGAACAGCTTGCAAGTTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62
 |||||
 Db 298 AACTTACAGAACAGCTGCGAAATTA-CGATTTGTACAAAACGATGCAACCTTCATCTTG 356
 |||||
 QY 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCCTTAACAGCTGGACCACA 122
 |||||
 Db 357 TTGATATCTGGAACATGATTAAGCTTTCCGAGACAATGGCCTTTATACACTGGACCATA 416
 |||||
 QY 123 GCACGGAGATCAGGGGTGTTCCCGCTGGAGACCGTCATCTCGTCCATCTACTATCATGTTG 182
 |||||
 Db 417 CCACCGAGATCA-GTGTGTCGCCCTCGAAACTGTCTCTCCATCTACTATCATGTTG 475
 |||||
 QY 183 AACACGGCCTTCCTCTCTACTCACCAGATCAGGGTGGACAGTCATCTAGTCTCTCTACTC 242
 |||||

Db 476 AACAGCGCCTTCCTCTACTCACCAAATAGTGTGGAACAATCTATCAGCCTCCTCCTC 535
 |||||
 QY 243 AATTTCATGTCGCCGCTACGACAGTGAGGGCGAGCCANGTTGACGGTGTGTTTCAGNT 302
 |||||
 Db 536 AACTTTATGATGTCATATGACAGTGAGGGCGAGCGAAGTTGACGGTATTTTCAGTT 595
 |||||
 QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 362
 |||||
 Db 596 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATATGTTTC 655
 |||||
 QY 363 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTNGGAAAGCTTGACCAGT 414
 |||||
 Db 656 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTTAGCAAGTTTGACCAGT 707
 |||||

RESULT 9
 AL524324 967 bp mRNA linear EST 13-FEB-2001
 LOCUS AL524324 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC006YE12 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL524324
 VERSION AL524324.1 GI:12787817
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 967)
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES Location/Qualifiers
 1..967
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DC006YE12"
 /clone_lib="LTI_NFL003_NBC3"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed
 by Life Technologies. Contact : Feng Liang Life
 Technologies, a division of Invitrogen 9800 Medical Center
 Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
 8371 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com"
 BASE COUNT 229 a 250 c 256 g 227 t 5 others
 ORIGIN
 Query Match 76.9%; Score 318.4; DB 9; Length 967;
 Best Local Similarity 88.6%; Pred. No. 1.8e-87;
 Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;
 QY 3 AACGTACAGAACAGCTTGCAAGTTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62
 |||||
 Db 322 AACTTACAGAACAGCTGCGAAATTA-CGATTTGTACAAAACGATGCAACCTTCATCTTG 380
 |||||
 QY 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCCTTAACAGCTGGACCACA 122
 |||||
 Db 381 TTGATATCTGGAACATGATTAAGCCTTTCCGAGACAATGGCCTTTATACACTGGACCATA 440
 |||||
 QY 123 GCACGGAGATCAGGGGTGTTCCCGCTGGAGACCGTCATCTCGTCCATCTACTATCATGTTG 182
 |||||
 Db 441 CCACCGAGATCA-GTGTGTCGCCCTCGAAACTGTCTCTCTCTACTATCATGTTG 499
 |||||

QY	183	AAACAGCGCCCTTCCTTCTACTCACCAGATCAGCGTGAGGAGTCCCATCAGTCTCCTACTC	242
Db	439	AAACAAGCGCCTTCCTTCTACTCACCAATTAGTGTGGAAACAATCTATCAGCCTCCTCCTC	498
QY	243	AATTTCATGGTGC GGCGCTTAGCACAGTGAAGCGCGGAGGANGTTGACCGTGTTCAGNT	302
Db	499	AAC TTTATGATGTG CATATACAGAGTGAAGCGCGGCAAGTTGACGGTATTTCAGTT	558
QY	303	AAAGCTATGTTAGCAACATGTGTGGTGGAAAAATGCTGGACAAAATTGAGATACATTTC	362
Db	559	AAAC TATGTTAGCAACATGTTGGTGGAAAAATGCTGGACAAAATTGAGATATGTTTC	618
QY	363	TCC CAGATGTCAGATTC CAATGGCTTAATGATGTTNGGAAAGCTTGACCACT	414
Db	619	TCCCAGATGTCAGATTC CAATGGCTTAATGATATATTAGCAAGTTTGACCACT	670
RESULT 11			
BI854854			
LOCUS			
DEFINITION BI854854 845 bp mRNA linear EST 10-OCT-2001			
mRNA sequence.			
ACCESSION BI854854			
VERSION BI854854.1 GI:15995601			
KEYWORDS			
SOURCE Est.			
ORGANISM house mouse			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.			
REFERENCE 1 (bases 1 to 845)			
NIH-MGC http://mgc.nci.nih.gov/.			
National Institutes of Health, Mammalian Gene Collection (MGC)			
Unpublished (1999)			
COMMENT Contact: Robert Strausberg, Ph.D.			
Email: cgabbs@mail.nih.gov			
Tissue Procurement: Jeffrey Green M.D.			
CDNA Library Preparation: Life Technologies, Inc.			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLML1987 row: j column: 07			
High quality sequence stop: 800.			
FEATURES			
source Location/Qualifiers			
1..845			
/organism="Mus musculus"			
/strain="FVB/N"			
/db_xref="taxon:10090"			
/clone="IMAGE:5387622"			
/clone_lib="NCI_CGAP_Mam6"			
/sex="female, virgin"			
/tissue_type="infiltrating ductal carcinoma"			
/dev_stage="5 months"			
/lab_host="DH10B"			
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;			
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.			
Library constructed by Life Technologies. Investigator			
providing samples: Jeffrey Green, M.D., NIH			
BASE COUNT 208 a 235 c 192 g 209 t 1 others			
ORIGIN			
Query Match 76.6%; Score 317; DB 13; Length 845;			
Best Local Similarity 98.5%; Pred. No. 4.6e-87;			
Matches 328; Conservative 0; Mismatches 4; Indels 1; Gaps 1;			
QY	82	CGAAGCTTTCCGAGACAATGGCCTTAACACGCTGGACCACAGCACGGAGATCAGGGGTCT	141
Db	1	CGAAGCTTTCCGAGACAATGGCCTTAACACGCTGGACCACAGCACGGAGATCA-GGCTGT	59
QY	142	NCCGCTGGAGACCGTCACTCTCGTCCATCTACTATCATGTTGAACAGGCCCTTCCTCTCA	201

JOURNAL
COMMENT

Unpublished (2001)
Other_ESTs: H3030G07-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdnaelgusun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgusun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3030 row: G column: 07
Seq primer: -21M13 Reverse
High quality sequence stop: 590
POLYA-No.

FEATURES
source

Location/Qualifiers
1. .590
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:H3030G07-5"
/db_xref="taxon:10090"
/clone="H3030G07"
/clone.lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 Kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."
Mol Genet 7: 1967-1978."
141 a 158 c 138 g 153 t

BASE COUNT
ORIGIN

Query Match 74.2%; Score 307; DB 12; Length 590;
Best Local Similarity 98.5%; Pred. No. 4.6e-84;
Matches 318; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 92 CGAGACATGGCCTTAACACGCTGGACACAGCAGGAGATCAGGCGTTCGCTGGA 151
DB 1 CGAGACATGGCCTTAACACGCTGGACACAGCAGGAGATCA-GGGTGTCCGCTGGA 59
QY 152 GACCGTCATCTCGTCCATCTACTATCAGTTGAACAGCGCCTTCCTTACTCACCAGAT 211
DB 60 GACCGTCATCTCGTCCATCTACTATCAGTTGAACAGCGCCTTCCTTACTCACCAGAT 119
QY 212 GACCGTGGAGCAGTCCATCAGTCTCCTACTCAATTTATGTCGCGCTCAGCAGTGA 271
DB 120 GACCGTGGAGCAGTCCATCAGTCTCCTACTCAATTTATGTCGCGCTCAGCAGTGA 179
QY 272 GGGCCGAGCAGTTGACCGGTGTTTTCAGTAAAGCTATGTTAGCAACCATGTGTGGTGG 331
DB 180 GGGCCGAGCAGTTGACCGGTGTTTTCAGTAAAGCTATGTTAGCAACCATGTGTGGTGG 239
QY 332 AAAAATGCTGGACAATTCAGATACATTTCTCCAGATGTCAGATTCATGCGCTTAAT 391
DB 240 AAAAATGCTGGACAATTCAGATACATTTCTCCAGATGTCAGATTCATGCGCTTAAT 299
QY 392 GATGTTNGAAAGCTTGGACCACT 414
||||| |||||||

Db 300 GATGTTTGGAAAGCTTGACCAGT 322

RESULT 14
BG805399
LOCUS

DEFINITION

0692-19 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,
mRNA sequence. 600 bp mRNA linear EST 20-DEC-2001

ACCESSION

BG805399

VERSION

BG805399.1 GI:17952332

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 600)

AUTHORS

Mu, X., Zhao, S., Pershad, R., Hsieh, T.-F., Scarpa, A., Wang, S.W.,

TITLE

White, R.A., Beremand, P.D., Thomas, T.L., Gan, L. and Klein, W.H.

JOURNAL

Gene expression in the developing mouse retina by EST sequencing

MEDLINE

and microarray analysis

COMMENT

Nucleic Acids Res. 29 (24), 4983-4993 (2001)

CONTACT

21671825

COMMENT

Contact: Klein WH

COMMENT

Department of Biochemistry and Molecular Biology

COMMENT

University of Texas M.D. Anderson Cancer Center

COMMENT

Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA

COMMENT

Tel: 713 792 3646

COMMENT

Fax: 713 790 0329

FEATURES

Location/Qualifiers

source

1. .600
/organism="Mus musculus"
/clone.lib="Mouse E14.5 retina lambda ZAP II Library"
/tissue_type="neural retina"
/dev_stage="embryonic day 14.5 post-fertilization"
/note="Vector: pAMP10 (Gibco); Cloned unidirectionally.
Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps
(Manniat); Cloning Technique: CUA Cloning (Clontechmp,
Life Technologies); Average insert size: 1.8 Kb;
Insertion site: TACGTCCACTGAATCTGAGTG---. Other
information regarding entire library may be found at
http://pga.swmed.edu/Data/Libraries/microarray_cdna_library.htm."

BASE COUNT 130 a 167 c 180 g 123 t

ORIGIN

Query Match 73.2%; Score 303; DB 12; Length 600;
Best Local Similarity 98.5%; Pred. No. 8.1e-83;
Matches 325; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 3 AACGTACAGACAGCTTCCAAAGTTACCGATTGTACAGAGCGGATGCAACCTTATCTTG 62
DB 273 AACGTACAGACAGCTTCCAAAGTTA-CGATTTGTACAGAGCGGATGCAACCTTATCTTG 331
QY 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCTTAACACGCTGGACACACA 122
DB 332 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCTTAACACGCTGGACACACA 391
QY 123 GCACGGAGATCAGCGGTGTCGCCCTGGAGACCGGTCTCTCGTCATCTACTATCATCTACTC 182
DB 392 GCACGGAGATCA-GCGTGTCCCGCTGGAGACCGGTCTCTCGTCATCTACTATCATCTACTC 450
QY 183 AACAGCGCCTTCTCTTACTCACCAGATCAGCGTGGAGAGCTCCATCAGTCTCTCTACTC 242
DB 451 AACAGCGCCTTCTCTTACTCACCAGATCAGCGTGGAGAGCTCCATCAGTCTCTCTACTC 510
QY 243 AATTTCATGTGCGCGCTACGACAGTGAAGCGCGGAGCANGTTGACCGTGTTCAGT 302
DB 511 AATTTCATGTGCGCGCTACGACAGTGAAGCGCGGAGCANGTTGACCGTGTTCAGT 570
QY 303 AAGCTATGTTAGCAACCATGTGTGGTGA 332
DB 571 AAGCTATGTTAGCAACCATGTGTGGTGA 600
||||| |||||||

RESULT 15
 BG469017
 LOCUS
 DEFINITION BG469017 830 bp mRNA linear EST 21-MAR-2001
 602510757p1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4644783 5',
 mRNA sequence.
 ACCESSION BG469017
 VERSION BG469017.1 GI:13401202
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 830)
 NIH-MGC http://mgc.nhl.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: NIH Intramural Sequencing Center
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LCM1417 row: b column: 16
 High quality sequence stop: 646.

FEATURES

source
 1..830
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4644783"
 /clone_lib="NIH_MGC_15"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5',
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)"
 BASE COUNT 213 a 207 c 221 g 189 t

Query Match 72.3%; Score 299.4; DB 12; Length 830;
 1st Local Similarity 87.2%; Pred. No. 1.3e-81;
 Matches 360; Conservative 0; Mismatches 50; Indels 3; Gaps 3;
 QY 3 AACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGGATGCAACCTTCATCTTG 62
 DB 265 AACTTACAGACAGCTGCAATTA-CGATTTGTACAAAACGATGCAACCTTCATCTTG 323
 QY 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAAATGGCCCTTAACACGCTGGACCACA 122
 DB 324 TTGATATCTGGAACATGATTGAAGCCCTCCGAGACAAATGGCCCTTAACACGCTGGACCACA 383
 QY 123 GCACGGAGATCAGCGGTGTCGCCCTGGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182
 DB 384 CCACCGAGATCA-GTGTGTCGCCCTCGAAACTGTCATCTCCTCCATCTACTATCAGTTG 442
 QY 183 AACAGCGCCTTCCTTCTACTACCAAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 242
 DB 443 AACAGCGCCTTCCTTCTACTACCAAAATAGTGTGGAACAATCTATCAGCCTCCTCCTC 502
 QY 243 AATTTTCATGTCGCCCTTACGACAGTGGAGGCCGAGGCAAGTTGACCGTGTTCAGNT 302
 DB 503 AACTTTATGATTCGTCATATGACATGAGGCGCGAGGCAAGTTGACCGTATTTTCAGTT 562
 QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 362

Db 563 AAAGCTATGTTAGCAACCATGTGTGGTGGAAACATGCTGGACAAATCGAGATATGTTATC 622
 QY 363 TCCAGATGTCAGATTCGAATGG-CTTAATGATGTTNGGAAAGCTTGACCACT 414
 Db 623 TCCAGATGTCAGATTCGAATGGCCTTCATGATATTCAGCAAAAGTTGACCACT 675
 Search completed: December 1, 2002, 17:42:10
 Job time : 1989 secs